# TEI P5:
# Guidelines for Electronic Text Encoding and Interchange

by the TEI Consortium

Originally edited by C.M. Sperberg-McQueen and Lou Burnard for the ACH-ALLC-ACL Text Encoding Initiative
Now entirely revised and expanded under the supervision of the Technical Council of the TEI Consortium

*The TEI Consortium*
*P5 Version 4.10.0. Last updated on 15th August 2025, revision ebf0834e1*

Text Encoding Initiative Consortium
2025

*The TEI Consortium*

***TEI P5:***
***Guidelines for Electronic Text Encoding and Interchange***

Text Encoding Initiative Consortium 2025

# The TEI Guidelines

The TEI Consortium

August 15, 2025

P5 Version 4.10.0. Last updated on 15th August 2025, revision ebf0834e1

Distributed under a Creative Commons Attribution 3.0 Unported License.
Copyright 2025 TEI Consortium.
All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

This software is provided by the copyright holders and contributors "as is" and any express or implied warranties, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose are disclaimed. In no event shall the copyright holder or contributors be liable for any direct, indirect, incidental, special, exemplary, or consequential damages (including, but not limited to, procurement of substitute goods or services; loss of use, data, or profits; or business interruption) however caused and on any theory of liability, whether in contract, strict liability, or tort (including negligence or otherwise) arising in any way out of the use of this software, even if advised of the possibility of such damage.

TEI material can be licensed differently depending on the use you intend to make of it. Hence it is made available under both the CC+BY and BSD-2 licences. The CC+BY licence is generally appropriate for usages which treat TEI content as data or documentation. The BSD-2 licence is generally appropriate for usage of TEI content in a software environment. For further information or clarification, please contact the TEI Consortium.

# Contents


i Releases of the TEI Guidelines . . . x
ii Dedication . . . xi
iii Preface and Acknowledgments . . . xii
iv About These Guidelines . . . xviii
iv.1 Structure and Notational Conventions of this Document . . . xviii
iv.2 Historical Background . . . xxv
iv.3 Future Developments and Version Numbers . . . xxvi
v A Gentle Introduction to XML . . . xxvii
v.1 What's Special about XML? . . . xxvii
v.2 Textual Structures . . . xxix
v.3 XML Structures . . . xxix
v.4 Validating a Document's Structure . . . xxxi
v.5 Complicating the Issue . . . xxxv
v.6 Attributes . . . xxxvi
v.7 Other Components of an XML Document . . . xxxix
v.8 Putting It All Together . . . xliii
vi Languages and Character Sets . . . xlvii
vi.1 Language Identification . . . xlvii
vi.2 Characters and Character Sets . . . xlviii
1 The TEI Infrastructure . . . 1
1.1 TEI Modules . . . 1
1.2 Defining a TEI Schema . . . 2
1.3 The TEI Class System . . . 4
1.4 Macros . . . 16
1.5 The TEI Infrastructure Module . . . 19
2 The TEI Header . . . 21
2.1 Organization of the TEI Header . . . 22
2.2 The File Description . . . 27
2.3 The Encoding Description . . . 39
2.4 The Profile Description . . . 57
2.5 Non-TEI Metadata . . . 65
2.6 The Revision Description . . . 66
2.7 Minimal and Recommended Headers . . . 67
2.8 Note for Library Cataloguers . . . 71
2.9 The TEI Header Module . . . 72
3 Elements Available in All TEI Documents . . . 73
3.1 Paragraphs . . . 74
3.2 Treatment of Punctuation . . . 75
3.3 Highlighting and Quotation . . . 78
3.4 Terms and Glosses, Ruby Annotations, and Equivalents and Descriptions . . . 87
3.5 Simple Editorial Changes . . . 95
3.6 Names, Numbers, Dates, Abbreviations, and Addresses . . . 103
3.7 Simple Links and Cross-References . . . 116
3.8 Lists . . . 119
3.9 Notes, Annotation, and Indexing . . . 123
3.10 Graphics and Other Non-textual Components . . . 130
3.11 Reference Systems . . . 131
3.12 Bibliographic Citations and References . . . 141
3.13 Passages of Verse or Drama . . . 167
3.14 Overview of the Core Module . . . 172
4 Default Text Structure . . . 174
4.1 Divisions of the Body . . . 176
4.2 Elements Common to All Divisions . . . 182

4.3 Grouped and Floating Texts . . . 189
4.4 Virtual Divisions . . . 198
4.5 Front Matter . . . 198
4.6 Title Pages . . . 201
4.7 Back Matter . . . 203
4.8 Module for Default Text Structure . . . 205
5 Characters, Glyphs, and Writing Modes . . . 206
5.1 Is Your Journey Really Necessary? . . . 206
5.2 Markup Constructs for Representation of Characters and Glyphs . . . 207
5.3 Annotating Characters . . . 212
5.4 Adding New Characters . . . 215
5.5 How to Use Code Points from the Private Use Area . . . 217
5.6 Writing Modes . . . 217
5.7 Examples of Different Writing Modes . . . 219
5.8 Text Rotation . . . 226
5.9 Caveat . . . 227
5.10 Formal Definition . . . 230
6 Verse . . . 231
6.1 Structural Divisions of Verse Texts . . . 231
6.2 Components of the Verse Line . . . 234
6.3 Encoding Textual Structures Across Verses . . . 237
6.4 Rhyme and Metrical Analysis . . . 237
6.5 Rhyme . . . 241
6.6 Metrical Notation Declaration . . . 243
6.7 Encoding Procedures for Other Verse Features . . . 246
6.8 Module for Verse . . . 246
7 Performance Texts . . . 247
7.1 Front and Back Matter . . . 247
7.2 The Body of a Performance Text . . . 255
7.3 Other Types of Performance Text . . . 269
7.4 Module for Performance Texts . . . 272
8 Transcriptions of Speech . . . 273
8.1 General Considerations and Overview . . . 273
8.2 Documenting the Source of Transcribed Speech . . . 275
8.3 Elements Unique to Spoken Texts . . . 277
8.4 Elements Defined Elsewhere . . . 285
8.5 Module for Transcribed Speech . . . 295
9 Computer-mediated Communication . . . 296
9.1 General Considerations . . . 296
9.2 Basic Units of CMC . . . 296
9.3 Encoding Unique to CMC . . . 298
9.4 CMC Macrostructure . . . 303
9.5 Documenting CMC (and providing general metadata) . . . 306
9.6 Recommendations for Encoding CMC Microstructure . . . 312
9.7 The TEI CMC Module . . . 318
10 Dictionaries . . . 319
10.1 Dictionary Body and Overall Structure . . . 319
10.2 The Structure of Dictionary Entries . . . 323
10.3 Top-level Constituents of Entries . . . 328
10.4 Headword and Pronunciation References . . . 346
10.5 Typographic and Lexical Information in Dictionary Data . . . 348
10.6 Unstructured Entries . . . 357
10.7 The Dictionary Module . . . 358
11 Manuscript Description . . . 359

11.1 Overview . . . 359
11.2 The Manuscript Description Element . . . 359
11.3 Phrase-level Elements . . . 363
11.4 The Manuscript Identifier . . . 372
11.5 The Manuscript Heading . . . 376
11.6 Intellectual Content . . . 376
11.7 Physical Description . . . 383
11.8 History . . . 395
11.9 Additional Information . . . 397
11.10 Manuscript Parts . . . 400
11.11 Manuscript Fragments . . . 401
11.12 Module for Manuscript Description . . . 402
12 Representation of Primary Sources . . . 403
12.1 Digital Facsimiles . . . 403
12.2 Combining Transcription with Facsimile . . . 411
12.3 Scope of Transcriptions . . . 419
12.4 Aspects of Layout . . . 454
12.5 Transcription and Ruby . . . 457
12.6 Headers, Footers, and Similar Matter . . . 457
12.7 Identifying Changes and Revisions . . . 457
12.8 Other Primary Source Features not Covered in these Guidelines . . . 461
12.9 Module for Transcription of Primary Sources . . . 461
13 Critical Apparatus . . . 462
13.1 The Apparatus Entry, Readings, and Witnesses . . . 463
13.2 Linking the Apparatus to the Text . . . 475
13.3 Using Apparatus Elements in Transcriptions . . . 481
13.4 Strategies for Encoding Variation . . . 483
13.5 Module for Critical Apparatus . . . 484
14 Names, Dates, People, and Places . . . 485
14.1 Attribute Classes Defined by This Module . . . 485
14.2 Names . . . 489
14.3 Biographical and Prosopographical Data . . . 501
14.4 Dates . . . 539
14.5 Module for Names and Dates . . . 544
15 Tables, Formulæ, Graphics, and Notated Music . . . 545
15.1 Tables . . . 545
15.2 Formulæ and Mathematical Expressions . . . 549
15.3 Notated Music in Written Text . . . 552
15.4 Specific Elements for Graphic Images . . . 555
15.5 Overview of Basic Graphics Concepts . . . 558
15.6 Graphic Image Formats . . . 558
15.7 Module for Tables, Formulæ, Notated Music, and Graphics . . . 560
16 Language Corpora . . . 562
16.1 Varieties of Composite Text . . . 562
16.2 Contextual Information . . . 566
16.3 Associating Contextual Information with a Text . . . 572
16.4 Linguistic Annotation of Corpora . . . 577
16.5 Recommendations for the Encoding of Large Corpora . . . 577
16.6 Module for Language Corpora . . . 578
17 Linking, Segmentation, and Alignment . . . 579
17.1 Links . . . 579
17.2 Pointing Mechanisms . . . 586
17.3 Blocks, Segments, and Anchors . . . 597
17.4 Synchronization . . . 603

17.5 Correspondence and Alignment . . . 607
17.6 Identical Elements and Virtual Copies . . . 614
17.7 Aggregation . . . 616
17.8 Alternation . . . 621
17.9 Stand-off Markup . . . 626
17.10 The standOff Container . . . 631
17.11 Annotations . . . 632
17.12 Connecting Analytic and Textual Markup . . . 634
17.13 Module for Linking, Segmentation, and Alignment . . . 634
18 Simple Analytic Mechanisms . . . 635
18.1 Linguistic Segment Categories . . . 635
18.2 Global Attributes for Simple Analyses . . . 644
18.3 Spans and Interpretations . . . 644
18.4 Linguistic Annotation . . . 649
18.5 Module for Analysis and Interpretation . . . 656
19 Feature Structures . . . 657
19.1 Organization of this Chapter . . . 657
19.2 Elementary Feature Structures and the Binary Feature Value . . . 657
19.3 Other Atomic Feature Values . . . 659
19.4 Feature Libraries and Feature-Value Libraries . . . 662
19.5 Feature Structures as Complex Feature Values . . . 664
19.6 Re-entrant Feature Structures . . . 666
19.7 Collections as Complex Feature Values . . . 667
19.8 Feature Value Expressions . . . 669
19.9 Default Values . . . 673
19.10 Linking Text and Analysis . . . 674
19.11 Feature System Declaration . . . 677
19.12 Formal Definition and Implementation . . . 690
20 Graphs, Networks, and Trees . . . 691
20.1 Graphs and Digraphs . . . 691
20.2 Trees . . . 703
20.3 Another Tree Notation . . . 709
20.4 Representing Textual Transmission . . . 716
20.5 Module for Graphs, Networks, and Trees . . . 718
21 Non-hierarchical Structures . . . 719
21.1 Multiple Encodings of the Same Information . . . 720
21.2 Boundary Marking with Empty Elements . . . 721
21.3 Fragmentation and Reconstitution of Virtual Elements . . . 724
21.4 Stand-off Markup . . . 727
21.5 Non-XML-based Approaches . . . 729
22 Certainty, Precision, and Responsibility . . . 730
22.1 Levels of Certainty . . . 730
22.2 Indications of Precision . . . 737
22.3 Attribution of Responsibility . . . 738
22.4 The Certainty Module . . . 739
23 Documentation Elements . . . 740
23.1 Phrase Level Documentary Elements . . . 741
23.2 Modules and Schemas . . . 742
23.3 Specification Elements . . . 744
23.4 Common Elements . . . 745
23.5 Element Specifications . . . 749
23.6 Class Specifications . . . 768
23.7 Macro Specifications . . . 770
23.8 Building a TEI Schema . . . 770

23.9 Module for Documentation Elements . . . 774
24 Using the TEI . . . 776
24.1 Serving TEI files with the TEI Media Type . . . 776
24.2 Obtaining the TEI . . . 776
24.3 Customization . . . 776
24.4 Conformance . . . 786
24.5 Implementation of an ODD System . . . 791
A Model Classes . . . 809
A.1 About the Model Classes Appendix . . . 809
B Attribute Classes . . . 832
B.1 About the Attribute Classes Appendix . . . 832
C Elements . . . 906
C.1 About the Elements Appendix . . . 906
D Attributes . . . 2450
D.1 About the Attributes Appendix . . . 2450
E Datatypes and Other Macros . . . 2456
E.1 About the Datatypes and Macros Appendix . . . 2456
F Bibliography . . . 2484
F.1 Works Cited in Examples in these Guidelines . . . 2484
F.2 Works Cited Elsewhere in the Text of these Guidelines . . . 2499
F.3 Reading List . . . 2502
G Deprecations . . . 2510
G.1 Items Scheduled to be Removed . . . 2510
H Prefatory Notes . . . 2511
H.1 Prefatory Note (November 2007) . . . 2511
H.2 Prefatory Note (March 2002) . . . 2511
H.3 Introductory Note (November 2001) . . . 2512
H.4 Introductory Note (June 2001) . . . 2512
H.5 Introductory Note (May 1999) . . . 2513
H.6 Preface (April 1994) . . . 2515
H.7 Acknowledgments . . . 2516
I Colophon . . . 2520

# i Releases of the TEI Guidelines

**P1** 1990, C.M. Sperberg-McQueen and Lou Burnard

**P2** 1992, C.M. Sperberg-McQueen and Lou Burnard

**P3** 1994, C.M. Sperberg-McQueen and Lou Burnard

**P4** 2001, Lou Burnard, Syd Bauman, and Steven DeRose

**P5** 2007, Lou Burnard and Syd Bauman

# ii Dedication

**In memoriam**

- Donald E. Walker
  22 November 1928 – 26 November 1993
- Antonio Zampolli
  1937 – 22 August 2003
- Sebastian Rahtz
  13 February 1955 – 15 March 2016
- C. Michael Sperberg-McQueen
  18 May 1954 – 16 August 2024

# iii Preface and Acknowledgments

At the time of this writing, eighteen years have passed since the TEI first released the fifth distinct version of the *Guidelines for Electronic Text Encoding and Interchange* in 2007. The release of the P5 Guidelines on 2 November 2007 was the happy result of six years of combined efforts of Editors Lou Burnard and Syd Bauman together with the recently established Technical Council, elected by the TEI community and charged with the ongoing maintenance and development of those Guidelines as an open-source project. While TEI P5 has been remarkably durable, it has survived and flourished through several infrastructure transitions and expanded in many new directions. The success of P5 can be seen in its incorporation of methods for standoff annotation; community activity supporting internationalization of glosses, descriptions, and examples; a new chapter on computer-mediated communication; and much experimental modeling that may be paving the way toward a P6, a topic of profound interest to Council members over the years.

These Guidelines have matured thanks to the stewardship of the Technical Council and the dedication of the Board over the years. The names and affiliations of all Board and Council members who served during the production of this edition of the Guidelines follow.

**Appointed Editors**

- 2001–2007: Syd Bauman (Brown Univesity)
- 2001–2007: Lou Burnard (University of Oxford)

**Board Chair**

- 2004–2005: Julia Flanders (Brown University)
- 2006: Matthew Zimmerman (New York University)
- 2007–2011: Daniel O'Donnell (University of Lethbridge)
- 2011: Martin Mueller (Northwestern University)
- 2011–2012: John Unsworth (University of Illinois at Urbana-Champaign / Brandeis University)
- 2012–2015: Elena Pierazzo (King's College London / Université Stendhal-Grenoble)
- 2016–2017: Michelle Dalmau (Indiana University)
- 2018–2021: Kathryn Tomasek (Wheaton College)
- 2022–2025: Diane K. Jakacki (Bucknell University)

**Elected Members of the Board of Directors**

- 2011–2013, 2015–2016: Marjorie Burghart (Centre National de la Recherche Scientifique / Ecole des Hautes Etudes en Sciences Sociales, Lyon)
- 2009–2014: Lou Burnard (Oxford University Computing Services / Independent Consultant)
- 2013–2014: Arianna Ciula (European Science Foundation, University of Roehampton)
- 2023–2025: Constance Crompton (University of Ottawa)
- 2020–2027: James Cummings (Newcastle University)
- 2011–2012: Marin Dacos (Revues.org / Center for Open Electronic Publishing, Cléo, Marseilles, France)
- 2015–2018: Michelle Dalmau (Indiana University)
- 2019–2020: Christiane Fritze (Austrian National Library)
- 2016: Kevin S. Hawkins (University of North Texas Libraries)
- 2021–2026: Diane K. Jakacki (Bucknell University)
- 2007–2008: John Lavagnino (King's College London)
- 2012–2013: Laura Mandell (Texas A & M University)
- 2022–2024: Wolfgang Meier (eXist Solutions)
- 2011–2013, 2015–2016: Martin Mueller (Northwestern University)
- 2017–2018: Kiyonori Nagasaki (International Institute for Digital Humanities, Tokyo)

- 2007–2011: Daniel O'Donnell (University of Lethbridge)
- 2021-2022: Ken Penner (St. Francis Xavier University)
- 2011-2015: Elena Pierazzo (King's College London / Université Stendhal-Grenoble)
- 2019–2026: Gimena del Rio Riande (Seminario de Edicion y Critica Textual, LINHD-UNED Madrid, University of Buenos Aires)
- 2017–2018: Laurent Romary (Inria)
- 2008–2011: Susan Schreibman (Trinity College Dublin)
- 2007–2008: Ray Siemens (University of Victoria)
- 2016–2021: Kathryn Tomasek (Wheaton College)
- 2025–2027: Magdalena Turska (Software Developer)
- 2016–2017: John Unsworth (University of Virginia)
- 2018–2019: Georg Vogeler (University of Graz)
- 2011–2014: John Walsh (Indiana University Bloomington)
- 2015–2020: Pip Willcox (University of Oxford / The National Archives UK)
- 2009–2010: Christian Wittern (Kyoto University)
- 2013–2014: Glen Worthey (Stanford University Libraries)

**Appointed (non-voting) members of the Board of Directors, in addition to the Technical Council Chair**

- Communications:
  - 2022–2023: Patricia O'Connor (University of Maynooth)
  - 2023–2025: Anna Sofia Lippolis (University of Bologna)
- Treasurer:
  - 2016–2025: Hugh Cayless (Duke University)
  - 2013–2015: John Unsworth (Brandeis University)
  - 2009–2013: Sarah Wells (University of Virginia)
  - 2007–2008: Daniel Pitti (University of Virginia)
- Website Editor:
  - 2024–2025: Hugh Cayless (Duke University)
  - 2015: Kevin S. Hawkins (University of North Texas Libraries)
  - 2019–2021: Luis Meneses (University of Victoria)
  - 2009–2014: David Sewell (University of Virginia Press)
- Partner Representatives:
  - 2011–2014: Sophie David (TGE Adonis)
  - 2011–2012: Wayne Graham (University of Virginia Library)
  - 2011–2014: Martin Wynne (Oxford Text Archive at University of Oxford)

**Technical Council Chair**

- 2002–2003: John Unsworth (University of Virginia)
- 2003–2007: Christian Wittern (Kyoto University)
- 2008–2011: Laurent Romary (Max Planck Digital Library / Inria / Humboldt-Universität zu Berlin)
- 2012–2014: James Cummings (University of Oxford)
- 2015–2017: Hugh Cayless (Duke University)
- 2018–2022: Martina Scholger (University of Graz)

- 2023–2025: Elisa Beshero-Bondar (Penn State Erie, The Behrend College)

**Members of the Technical Council Appointed by the TEI Board**

- 2001–2009: Sebastian Rahtz (University of Oxford)
- 2010–2014: Lou Burnard (University of Oxford / TGE Adonis / Independent Consultant)

**Elected Members of the Technical Council**

- 2011–2012: Piotr Bański (University of Warsaw)
- 2010–2013: Brett Barney (University of Nebraska)
- 2013–2026: Syd Bauman (Brown University / Northeastern University)
- 2021–2025: Helena Bermúdez Sabel (Université de Neuchâtel / JinnTec)
- 2016–2027: Elisa Beshero-Bondar (University of Pittsburgh at Greensburg / Penn State Erie, The Behrend College)
- 2022–2025: Elli Bleeker (Huygens Institute for the History of the Netherlands)
- 2003–2006, 2017–2018: Alejandro Bia (University of Alicante)
- 2019–2020: Vanessa Bigot Juloux (Ecole Pratique des Hautes Etudes / Paris Sciences et Lettres / Andrews University, Michigan)
- 2004–2007: David Birnbaum (University of Pittsburgh)
- 2008–2013: Gabriel Bodard (King's College London)
- 2008–2009: Peter Boot (Huygens Institute for Netherlands History)
- 2020–2021: Meaghan Brown (Papers of the Bibliographical Society of America)
- 2007–2008: Tone Merete Bruvik (University of Bergen)
- 2015–2016: Lou Burnard (Independent Consultant)
- 2013–2023: Hugh Cayless (New York University / Duke University)
- 2014–2015: Fabio Ciotti (University of Rome "Tor Vergata")
- 2007–2009: Arianna Ciula (King's College London / European Science Foundation)
- 2005–2019: James Cummings (University of Oxford / Newcastle University)
- 2020: Nicholas Cole (Pembroke College Oxford)
- 2002–2007, 2010: Matthew Driscoll (University of Copenhagen)
- 2002–2004: David Durand (Ingenta plc)
- 2002–2004: Tomas Erjavec (Jozef Stefan Institute, Ljubljana)
- 2024: Gustavo Fernández Riva (University Library of Heidelberg)
- 2014–2015: Stefanie Gehrke (Biblissima)
- 2010–2013: Kevin Hawkins (University of Michigan)
- 2025–2027: Ulrike Henny-Krahmer (University of Rostock)
- 2010–2015, 2025–2027: Martin Holmes (University of Victoria)
- 2002: Fotis Jannidis (University of Munich)
- 2021–2023: Janelle Jenstad (University of Victoria)
- 2006: Amit Kumar (University of Illinois at Urbana-Champaign)
- 2020–2021: Jessica H. Lu (University of Maryland)
- 2015–2016: Stefan Majewski (Austrian National Library)
- 2002: Martin Mueller (Northwestern University)
- 2013–2014, 2016–2019: Elli Mylonas (Brown University)
- 2010–2011: Julianne Nyhan (University of Trier / University College London)

- 2023–2025: Patricia O'Connor (Independent Researcher / University of Maynooth)
- 2008–2011: Elena Pierazzo (King's College London)
- 2006–2007, 2009–2010: Dot Porter (University of Kentucky / Digital Humanities Observatory / Indiana University)
- 2002–2003: Merillee Proffitt (Research Libraries Group)
- 2002: Peter Robinson (De Montfort University)
- 2009–2014: Sebastian Rahtz (University of Oxford)
- 2002: Geoffrey Rockwell (Macmaster University)
- 2024–2026: Torsten Roeder (University of Würzburg)
- 2002–2007: Laurent Romary (Inria / CNRS / Max Planck Digital Library)
- 2008–2009, 2012–2015: Paul Schaffner (University of Michigan)
- 2016–2027: Martina Scholger (University of Graz)
- 2003–2007: Susan Schreibman (University of Maryland)
- 2022–2024: Sabine Seifert (University of Potsdam)
- 2008–2009: David Sewell (University of Virginia)
- 2004–2005: Natasha Smith (University of North Carolina at Chapel Hill)
- 2014–2022: Peter Stadler (Carl-Maria-von-Weber-Gesamtausgabe / University of Paderborn)
- 2017–2019: Sarah Stanley (Florida State University)
- 2023–2026: Joey Takeda (Digital Humanities Innovation Lab, Simon Fraser University)
- 2008–2009: Manfred Thaller (University of Cologne)
- 2006–2007: Conal Tuohy (Victoria University of Wellington)
- 2016–2024: Magdalena Turska (eXist Solutions / University of Oxford)
- 2004–2005: Edward Vanhoutte (Royal Academy of Dutch Language and Literature)
- 2015–2026: Raffaele Viglianti (University of Maryland)
- 2005–2008: John Walsh (Indiana University)
- 2012–2013: Rebecca Welzenbach (University of Michigan)
- 2002–2005: Perry Willett (Indiana University / University of Michigan)
- 2011–2012: Stuart Yeates (New Zealand Electronic Text Centre)

The bulk of the Council's work has been carried out by email and by regular teleconferences and virtual work sessions. In addition, the Council has held many two-day face-to-face meetings. During development, production, and maintenance of P5, these meetings were generously hosted in person and online by many institutions, as documented:

- 2002: King's College, London
- 2003: Oxford University Computing Services
- 2004: Royal Academy of Dutch Language and Literature, Ghent
- 2005: Association française de normalisation, Paris
- 2006: Institute for Research in Humanities, Kyoto University
- 2007: Berlin-Brandenburgische Akademie der Wissenschaften, Berlin
- 2008: National University of Ireland, Galway
- 2009: Ecole Normale Supérieure, Lyon
- 2010: Royal Irish Academy, Dublin
- 2011: Big Ten Center, Chicago

- 2011: Inria, Paris
- 2012: University of Michigan, Ann Arbor
- 2012: University of Oxford
- 2013: Brown University
- 2013: University of Oxford
- 2014: University of Oxford
- 2014: Duke University
- 2015: University of Michigan, Ann Arbor
- 2015: Lyon
- 2016: Brown University
- 2016: Austrian Academy of Sciences
- 2017: Prague
- 2017: Victoria
- 2018: Cologne Center for eHumanities, Cologne University
- 2018: University of Tokyo
- 2019: Folger Shakespeare Library, Washington, D.C.
- 2019: Centre for Information Modeling, University of Graz
- 2022: Newcastle University
- 2023: University of Guelph
- 2023: Paderborn University
- 2024: Universidad del Salvador, Buenos Aires

During the production of TEI P5, the Council chartered a number of smaller workgroups and similar activities, each of which made significant contribution to the intellectual content of the work. Active members of these are listed below:

**Character Set Workgroup** Active between July 2001 and January 2005, this group revised and developed the recommendations now forming chapters *vi Languages and Character Sets* and *5. Characters, Glyphs, and Writing Modes*. It was chaired by Christian Wittern, and its membership included: Deborah Anderson (Berkeley); Michael Beddow (independent scholar); David Birnbaum (University of Pittsburgh); Martin Duerst (W3C/Keio University); Patrick Durusau (Society of Biblical Literature); Tomohiko Morioka (Kyoto University); and Espen Ore (National Library of Norway).

**Meta Taskforce** Active between February 2003 and February 2005, this group developed the material now forming *23. Documentation Elements*. It was chaired by Sebastian Rahtz, and its membership included: Alejandro Bia; David G. Durand; Laurent Romary; Norman Walsh (Sun Microsystems); and Christian Wittern.

**Workgroup on Stand-Off Markup, XLink and XPointer** Active between February 2002 and January 2006, this group reviewed and expanded the material now largely forming part of *17. Linking, Segmentation, and Alignment*. It was chaired by David G. Durand, and its membership included: Jean Carletta (Edinburgh University); Chris Caton (University of Oxford); Jessica P. Hekman (Ingenta plc); Nancy M. Ide (Vassar College); and Fabio Vitali (University of Bologna).

**Manuscript Description Task Force** Active between February 2003 and December 2005, this group reviewed and finalised the material now forming *11. Manuscript Description*. It was chaired by Matthew Driscoll and comprised David Birnbaum and Merrilee Proffitt, in addition to the TEI Editors.

**Names and Places Activity** Active between January 2006 and May 2007, this group formulated the new material now forming part of *14. Names, Dates, People, and Places*. It was chaired by Matthew Driscoll. and its membership included Gabriel Bodard (King's College London); Arianna Ciula; James Cummings; Tom Elliott (University of North Carolina at Chapel Hill); Øyvind Eide (University of Oslo); Leif Isaksen (Oxford

Archaeology plc); Richard Light (private consultant); Tadeusz Piotrowski (Opole University); Sebastian Rahtz; and Tatiana Timcenko (Vilnius University).

**Joint TEI/ISO Activity on Feature Structures** Active between January 2003 and August 2007, this group reviewed the material now presented in *19. Feature Structures* and revised it for inclusion in ISO Standard 24610. It was chaired by Kiyong Lee (Korea University), and its active membership included the following: Harry Bunt (Tilburg); Lionel Clément (INRIA); Eric de la Clergerie (INRIA); Thierry Declerck (Saarbrücken); Patrick Drouin (University of Montréal); Lee Gillam (Surrey University); and Kōiti Hasida (ICOT).

From 2000 to 2008 the TEI had two appointed Editors, Lou Burnard (University of Oxford) and Syd Bauman (Brown University), who served *ex officio* on the Council and, as far as possible, on all Council workgroups.

The Council also oversees an Internationalization and Localization project, led by Sebastian Rahtz and with funding from the ALLC. This activity, ongoing since October 2005, is engaged in translating key parts of the P5 source into a variety of languages.

Production of the translations currently included in P5 has been coordinated by the following:

**Chinese** Marcus Bingenheimer (Chung-hwa Institute of Buddhist Studies, Taipei / Temple University) and Weining Hwang (Würzburg University)

**French** Pierre-Yves Duchemin (ENSSIB); Jean-Luc Benoit (ATILF); Anila Angjeli (BnF); Joëlle Bellec Martini (BnF); Marie-France Claerebout (Aldine); Magali Le Coënt (BIUSJ); Florence Clavaud (EnC); Cécile Pierre (BIUSJ).

**German** Werner Wegstein (Würzburg University), Franz Fischer (Cologne University), Martina Scholger (University of Graz) and Peter Stadler (Paderborn University)

**Italian** Marco Venuti (University of Venice) and Letizia Cirillo (University of Bologna)

**Japanese** Ohya Kazushi (Tsurumi University), Kiyonori Nagasaki (University of Tokyo) and Martin Holmes (University of Victoria)

**Korean** Beom-mo Kang and Jungha Hong (Korea University)

**Spanish** Carmen Arronis Llopis (University of Alicante), Alejandro Bia (Miguel Hernández University), Gimena del Rio Riande (Consejo Nacional de Investigaciones Científicas y Técnicas) and Susanna Allés-Torrent (University of Miami)

Any one who works closely with the TEI Guidelines, whether as translator, editor, or reader is constantly reminded of the ambitious scope and exceptionally high editorial standards set by the original project, now over a third of a century ago. It is appropriate therefore to retain a sense of the history of this document, as it has evolved since its first appearance in 1990, and to acknowledge with gratitude the contributions made to that evolution by very many individuals and institutions around the world. The original prefatory notes to each major edition of the Guidelines recording these names are therefore preserved in an appendix to the current edition (see *Appendix H Prefatory Notes*).

# iv About These Guidelines

These Guidelines have been developed and are maintained by the Text Encoding Initiative Consortium (TEI); see *iv.2 Historical Background*. They are addressed to anyone who works with any kind of textual resource in digital form.

They make recommendations about suitable ways of representing those features of textual resources which need to be identified explicitly in order to facilitate processing by computer programs. In particular, they specify a set of markers (or *tags*) which may be inserted in the electronic representation of the text, in order to mark the text structure and other features of interest. Many, or most, computer programs depend on the presence of such explicit markers for their functionality, since without them a digitized text appears to be nothing but a sequence of undifferentiated bits. The success of the World Wide Web, for example, is partly a consequence of its use of such markup to indicate such features as headings and lists on individual pages, and to indicate links between pages. The process of inserting such explicit markers for implicit textual features is often called 'markup', or equivalently within this work 'encoding'; the term 'tagging' is also used informally. We use the term *encoding scheme* or *markup language* to denote the complete set of rules associated with the use of markup in a given context; we use the term *markup vocabulary* for the specific set of markers or named distinctions employed by a given encoding scheme. Thus, this work both describes the TEI encoding scheme, and documents the TEI markup vocabulary.

The TEI encoding scheme is of particular usefulness in facilitating the loss-free interchange of data amongst individuals and research groups using different programs, computer systems, or application software. Since they contain an inventory of the features most often deployed for computer-based text processing, these Guidelines are also useful as a starting point for those designing new systems and creating new materials, even where interchange of information is not a primary objective.

These Guidelines apply to texts in any natural language, of any date, in any literary genre or text type, without restriction on form or content. They treat both continuous materials ('running text') and discontinuous materials such as dictionaries and linguistic corpora. Though principally directed to the needs of the scholarly research community, these Guidelines are not restricted to esoteric academic applications. They are also useful for librarians maintaining and documenting electronic materials, and for publishers and others creating or distributing electronic texts. Although they focus on problems of representing in electronic form texts which already exist in traditional media, these Guidelines are also applicable to textual material which is 'born digital'. We believe them to be adequate to the widest variety of currently existing practices in using digital textual data, but by no means limited to them.

The rules and recommendations made in these Guidelines are expressed in terms of what is currently the most widely-used markup language for digital resources of all kinds: the Extensible Markup Language (XML), as defined by the World Wide Web Consortium's XML Recommendation. However, the TEI encoding scheme itself does not depend on this language; it was originally formulated in terms of SGML (the ISO Standard Generalized Markup Language), a predecessor of XML, and may in future years be re-expressed in other ways as the field of markup develops and matures. For more information on markup languages see chapter *v A Gentle Introduction to XML*; for more information on the associated character encoding issues see chapter *vi Languages and Character Sets*.

This document provides the authoritative and complete statement of the requirements and usage of the TEI encoding scheme. As such, although it includes numerous small examples, it must be stressed that this work is intended to be a reference manual rather than a tutorial guide.

The remainder of this chapter comprises three sections. The first gives an overview of the structure and notational conventions used throughout these Guidelines. The second enumerates the design principles underlying the TEI scheme and the application environments in which it may be found useful. Finally, the third section gives a brief account of the origins and development of the Text Encoding Initiative itself.

## iv.1 Structure and Notational Conventions of this Document

The remaining two sections of the front matter to these Guidelines provide background tutorial material for those unfamiliar with basic markup technologies. Following the present introductory section, we present a detailed introduction to XML itself, intended to cover in a relatively painless manner as much as the novice user of the TEI scheme needs to know about markup languages in general and XML in particular. This is followed by a discussion of the general principles underlying current practice in the representation of different languages and writing systems in digital form. This chapter is largely intended for the user unfamiliar with the Unicode encoding systems, though the expert may also find its historical overview of interest.

The body of this edition of these Guidelines proper contains 23 chapters arranged in increasing order of specialist interest. The first five chapters discuss in depth matters likely to be of importance to anyone intending to apply the

TEI scheme to virtually any kind of text. The next seven focus on particular kinds of text: verse, drama, spoken text, dictionaries, and manuscript materials. The next nine chapters deal with a wide range of topics, one or more of which are likely to be of interest in specialist applications of various kinds. The last two chapters deal with the XML encoding used to represent the TEI scheme itself, and provide technical information about its implementation. The last chapter also defines the notion of TEI conformance and its implications for interchange of materials produced according to these Guidelines.

As noted above, this is a reference work, and is not intended to be read through from beginning to end. However, the reader wishing to understand the full potential of the TEI scheme will need a thorough grasp of the material covered by the first four chapters and the last two. Beyond that, the reader is recommended to select according to their specific interests: one of the strengths of the TEI architecture is its modular nature.

As far as possible, extensive cross referencing is provided wherever related topics are dealt with; these are particularly effective in the online version of these Guidelines. In addition, a series of technical appendixes provide detailed formal definitions for every element, every class, and every macro discussed in the body of the work; these are also cross linked as appropriate. Finally, a detailed bibliography is provided, which identifies the source of many examples cited in the text as well as documenting works referred to, and listing other relevant publications.

As an aid to the reader, most chapters of these Guidelines follow the same basic organization. The chapter begins with an overview of the subjects treated within it, linked to the following subsections. Within each section where new elements are described, a summary table is first given, which provides their names and a brief description of their intended usage. This is then followed where appropriate by further discussion of each element, including wherever possible usage examples taken somewhat eclectically from a variety of real sources. These examples are not intended to be exhaustive, but rather to suggest typical ways in which the elements concerned may usefully be applied. Where appropriate, a link to a statement of the source for most examples is provided in the online version. Within the examples, use of whitespace such as newlines or indentation is simply intended to aid legibility, and is not prescriptive or normative.

Wherever TEI elements or classes are mentioned in the text, they are linked in the online version to the relevant reference specification for the element or class concerned. Element names are always given in the form <name>, where 'name' is the *generic identifier* of the element; empty elements such as <pb> or <anchor> include a closing slash to distinguish them wherever they are discussed. References to attributes take the form *attname*, where 'attname' is the name of the attribute. References to classes are also presented as links, for example `model.divLike` for a model class, and `att.global` for an attribute class.

### iv.1.1 The use of modal verbs

In general the TEI guidelines try to be careful when using modal verbs and phrases such as 'must', 'must not, 'should', 'should not' and 'may'. In terms of the meanings, these generally follow `https://tools.ietf.org/html/bcp14` in the different meanings of these words. In particular:

**MUST** This word, or the terms "REQUIRED" or "SHALL", means that this is an absolute requirement of the TEI Guidelines for production of a TEI conformant file.

**MUST NOT** This phrase, or the phrase "SHALL NOT", means that this is an absolute prohibition of the TEI Guidelines for production of a TEI conformant file.

**SHOULD** This word, or the adjective "RECOMMENDED", means that there may exist valid reasons in particular circumstances to ignore a particular recommendation, but the full implications must be understood and carefully weighed before choosing a different course.

**SHOULD NOT** This phrase, or the phrase "NOT RECOMMENDED" means that there may exist valid reasons in particular circumstances when the particular behavior is acceptable or even useful, but the full implications should be understood and the case carefully weighed before implementing any behavior so described.

**MAY** This word, or the adjective "OPTIONAL", means that a recommendation is truly optional. One user may choose to follow the recommendation because a particular project requires it or feels that it enhances their work while another project may choose to not follow this recommendation.

However, the prose of the TEI Guidelines continually undergoes revision and all uses may not reflect this. Please report any errors or unclear use of these words.

## iv.1.2 TEI Naming Conventions

These Guidelines use a more or less consistent set of conventions in the naming of XML elements and classes. This section summarizes those conventions.

### Element and Attribute Names

An unadorned name such as 'blort' is the name of a TEI element or attribute. [1].

The following conventions apply to the choice of names:

- Elements are given generic identifiers as far as possible consisting of one or more *tokens*, by which we mean whole words or recognisable abbreviations of them, taken from the English language.
- Where an element name contains more than one token, the first letter of the second token, and of any subsequent ones, is capitalized, as in for example <biblStruct>, <listPerson>. This 'camelCasing' is used also for attribute names and symbolic values.
- Module names also use whole words, for the most part, but are always all lower case.
- The specification for an element or attribute whose name contains abbreviations generally also includes a <gloss> element providing the expanded sense of the name.
- An element specification may also contain approved translations for element or attribute names in one or more other languages using the <altIdent> element; this is not however generally done in TEI P5.

Whole words are generally preferred for clarity. The following abbreviations are however commonly used within generic identifiers:

**att** attribute

**bibl** bibliographic description or reference in a bibliography

**cat** category, especially as used in text classification

**char** character, typically a Unicode character

**doc** document: this usually refers to the original source document which is being encoded,

**decl** declaration: has a specific sense in the TEI header, as discussed in *2.1.2. Types of Content in the TEI Header*

**desc** description: has a specific sense in the TEI header, as discussed in *2.1.2. Types of Content in the TEI Header*

**grp** group. In TEI usage, a group is distinguished from a list in that the former associates several objects which act as a single entity, while the latter does not. For example, a <linkGrp> combines several <link> elements which have certain properties in common, whereas a <listBibl> simply lists a number of otherwise unrelated <bibl> elements.

**interp** interpretation or analysis

**lang** (natural) language

**ms** manuscript

**org** organization, that is, a named group of people or legal entity

**rdg** reading or version found in a specific witness

**ref** reference or link

**spec** technical specification or definition

**stmt** statement: used in a specific sense in the TEI header, as discussed in *2.1.2. Types of Content in the TEI Header*

**struct** structured: that is, containing a specific set of named elements rather than 'mixed content'

**val** value, for example of a variable or an attribute

**wit** witness: that is, a specific document which attests specific readings in a textual tradition or apparatus

[1] During generation of TEI RELAX NG schema fragments, the patterns corresponding with these TEI names are given a prefix `tei` to allow them to co-exist with names from other XML namespace. This prefix is not visible to the end user, and is not used in TEI documentation. When generating multi-namespace schemas, however, the user needs to be aware of them.

Some abbreviations are used inconsistently: for example, <add> is an addition, and <addSpan> is a spanning addition, but <addName> is an additional name, not the name of an addition. Such inconsistencies are relatively few in number, and it is hoped to remove them in subsequent revisions of these Guidelines.

Some elements have very short abbreviated names: these are for the most part elements which are likely to be used very frequently in a marked up text, for example <p> (paragraph), <s> (segment) <hi> (highlighted phrase), <ptr> (pointer),  (division) etc. We do not specifically list such elements here: as noted above, an expansion of each such abbreviated name is provided within the documentation using the <gloss> element .

### iv.1.3 Class, Macro, and Datatype Names

All named objects other than elements and attributes have one of the following prefixes, which indicate whether the object is a module, an attribute class, a model class, a datatype, or a macro:

| Component | Name | Example |
|---|---|---|
| Attribute Classes | att.* | att.global |
| Model Classes | model.* | model.biblPart |
| Macros | macro.* | macro.paraContent |
| Datatypes | teidata.* | teidata.pointer |

The concepts of model class, attribute class, etc. are defined in *1. The TEI Infrastructure*. Here we simply note some conventions about their naming.

The following rules apply to attribute class names:

- Attribute class names take the form `att.xxx`, where `xxx` is typically an adjective, or a series of adjectives separated by dots, describing a property common to the attributes which make up the class.
- Attributes with the same name are considered to have the same semantics, whether the attribute is inherited from a class, or locally defined.

The following rules apply to model class names:

- Model classes have names beginning `model.` followed by a *root name*, and zero or more suffixes as described below.
- A root name may be the name of an element, generally the prototypical parent or sibling for elements which are members of the class.
- The first suffix should be `Part`, if the class members are all children of the element named rootname; or `Like`, if the class members are all siblings of the element named `rootname`.
- The rootname `global` is used to indicate that class members are permitted anywhere in a TEI document.
- Additional suffixes may be added, prefixed by a dot, to distinguish subclasses, semantic or structural.

For example, the class of elements which can form part of a  is called `model.divPart`. This class includes as a subclass the elements which can form part of a  in a spoken text, which is named `model.divPart.spoken`

### iv.1.4 Design Principles

Because of its roots in the humanities research community, the TEI scheme is driven by its original goal of serving the needs of research, and is therefore committed to providing a maximum of comprehensibility, flexibility, and extensibility. More specific design goals of the TEI have been that these Guidelines should:

- provide a standard format for data interchange
- provide guidance for the encoding of texts in this format
- support the encoding of all kinds of features of all kinds of texts studied by researchers
- be application independent

This has led to a number of important design decisions, such as:

- the choice of XML and Unicode
- the provision of a large predefined tag set
- encodings for different views of text

- alternative encodings for the same textual features
- mechanisms for user-defined modification of the scheme

We discuss some of these goals in more detail below.

The goal of creating a common interchange format which is application independent requires the definition of a specific markup syntax as well as the definition of a large set of elements or concepts. The syntax of the recommendations made in this document conforms to the World Wide Web Consortium's XML Recommendation (**?**) but their definition is as far as possible independent of any particular schema language.

The goal of providing guidance for text encoding suggests that recommendations be made as to what textual features should be recorded in various situations. However, when selecting certain features for encoding in preference to others, these Guidelines have tended to prefer generic solutions to specific ones, and to avoid areas where no consensus exists, while attempting to accommodate as many diverse views as feasible. Consequently, the TEI Guidelines make (with relatively rare exceptions) no suggestions or restrictions as to the relative importance of textual features. The philosophy of these Guidelines is 'if you want to encode this feature, do it this way'—but very few features are mandatory. In the same spirit, while these Guidelines very rarely require you to encode any particular feature, they do require you to be honest about which features you have encoded, that is, to respect the meanings and usage rules they recommend for specific elements and attributes proposed.

The requirement to support all kinds of materials likely to be of interest in research has largely conditioned the development of the TEI into a very flexible and modular system. The development of other XML vocabularies or standards is typically motivated by the desire to create a single fully specified encoding scheme for use in a well-defined application domain. By contrast, the TEI is intended for use in a large number of rather ill-defined and often overlapping domains. It achieves its generality by means of the modular architecture described in *1. The TEI Infrastructure* which enables each user to create a schema appropriate to their needs without compromising the interoperability of their data.

The Guidelines have been written largely with a focus on text capture (i.e. the representation in electronic form of an already existing copy text in another medium) rather than text creation (where no such copy text exists). Hence the frequent use of terms like 'transcription', 'original', 'copy text', etc. However, these Guidelines are equally applicable to text creation, although certain elements, such as <sourceDesc>, and certain attributes, such as the rendition indicators, will not be relevant in this case.

Concerning text capture the TEI Guidelines do not specify a particular approach to the problem of fidelity to the source text and recoverability of the original; such a choice is the responsibility of the text encoder. The current version of these Guidelines, however, provides a more fully elaborated set of tags for markup of rhetorical, linguistic, and simple typographic characteristics of the text than for detailed markup of page layout or for fine distinctions among type fonts or manuscript hands. It should be noted also that, with the present version of these Guidelines, it is no longer necessarily the case that an unmediated version of the source text can be recovered from an encoded text simply by removing the markup.

In these Guidelines, no hard and fast distinction is drawn between 'objective' and 'subjective' information or between 'representation' and 'interpretation'. These distinctions, though widely made and often useful in narrow, well-defined contexts, are perhaps best interpreted as distinctions between issues on which there is a scholarly consensus and issues where no such consensus exists. Such consensus has been, and no doubt will be, subject to change. The TEI Guidelines do not make suggestions or restrictions as to which of these features should be encoded. The use of the terms *descriptive* and *interpretive* about different types of encoding in these Guidelines is not intended to support any particular view on these theoretical issues. Historically, it reflects a purely practical division of responsibility amongst the original working committees (see further *iv.2 Historical Background*).

In general, the accuracy and the reliability of the encoding and the appropriateness of the interpretation is for the individual user of the text to determine. The Guidelines provide a means of documenting the encoding in such a way that a user of the text can know the reasoning behind that encoding, and the general interpretive decisions on which it is based. The TEI header should be used to document and justify many such aspects of the encoding, but the choice of TEI elements for a particular feature is in itself a statement about the interpretation reached by the encoder.

In many situations more than one view of a text is needed since no absolute recommendation to embody one specific view of text can apply to all texts and all approaches to them. Within limits, the syntax of XML ensures that some encodings can be ignored for some purposes. To enable encoding multiple views, these Guidelines not only treat a variety of textual features, but sometimes provide several alternative encodings for what appear to

be identical textual phenomena. These Guidelines offer the possibility of encoding many different views of the text, simultaneously if necessary. Where different views of the formal structure of a text are required, as opposed to different annotations on a single structural view, however, the formal syntax of XML (which requires a single hierarchical view of text structure) poses some problems; recommendations concerning ways of overcoming or circumventing that restriction are discussed in chapter *21. Non-hierarchical Structures*.

In brief, the TEI Guidelines define a general-purpose encoding scheme which makes it possible to encode different views of text, possibly intended for different applications, serving the majority of scholarly purposes of text studies in the humanities. Because no predefined encoding scheme can possibly serve all research purposes, the TEI scheme is designed to facilitate both selection from a wide range of predefined markup choices, and the addition of new (non-TEI) markup options. By providing a formally verifiable means of extending the TEI recommendations, the TEI makes it simple for such user-identified modifications to be incorporated into future releases of these Guidelines as they evolve. The underlying mechanisms which support these aspects of the scheme are introduced in chapter *1. The TEI Infrastructure*, and detailed discussions of their use provided in chapter *24. Using the TEI*.

## iv.1.5 Intended Use

We envisage three primary functions for these Guidelines:

- guidance for individual or local practice in text creation and data capture;
- support of data interchange;
- support of application-independent local processing.

These three functions are so thoroughly interwoven in practice that it is hardly possible to address any one without addressing the others. However, the distinction provides a useful framework for discussing the possible role of these Guidelines in work with electronic texts.

### Use in Text Capture and Text Creation

The description of textual features found in the chapters which follow should provide a useful checklist from which scholars planning to create electronic texts should select the subset of features suitable for their project.

Problems specific to text creation or text 'capture' have not been considered explicitly in this document. These Guidelines are not concerned with the process by which a digital text comes into being: it can be typed by hand, scanned from a printed book or typescript, read from a typesetter's tape, or acquired from another researcher who may have used another markup scheme (or no explicit markup at all).

We include here only some general points which are often raised about markup and the process of data capture.

XML can appear distressingly verbose, particularly when (as in these Guidelines) the names of tags and attributes are chosen for clarity and not for brevity. Editor macros and keyboard shortcuts can allow a typist to enter frequently used tags with single keystrokes. It is often possible to transform word-processed or scanned text automatically. Markup-aware software can help with maintaining the hierarchical structure of the document, and display the document with visual formatting rather than raw tags.

The techniques described in chapter *24.3. Customization* may be used to develop simpler data capture TEI-conformant schemas, for example with limited numbers of elements, or with shorter names for the tags being used most often. Documents created with such schemas may then be automatically converted to a more elaborated TEI form.

### Use for Interchange

The TEI format may simply be used as an interchange format, permitting projects to share resources even when their local encoding schemes differ. If there are n different encoding formats, to provide mappings between each possible pair of formats requires n×(n-1) translations; with an interchange format, only 2×n such mappings are needed. However, for such translations to be carried out without loss of information, the interchange format chosen must be as expressive (in a formal sense) as any of the target formats; this is a further reason for the TEI's provision of both highly abstract or generic encodings and highly specific ones.

To translate between any pair of encoding schemes implies:

1. identifying the sets of textual features distinguished by the two schemes;
2. determining where the two sets of features correspond;
3. creating a suitable set of mappings.

For example, to translate from encoding scheme X into the TEI scheme:

1. Make a list of all the textual features distinguished in X.
2. Identify the corresponding feature in the TEI scheme. There are three possibilities for each feature:

   (a) the feature exists in both X and the TEI scheme;

   (b) X has a feature which is absent from the TEI scheme;

   (c) X has a feature which corresponds with more than one feature in the TEI scheme.

   The first case is a trivial renaming. The second will require an extension to the TEI scheme, as described in chapter *24.3. Customization*. The third is more problematic, but not impossible, provided that a consistent choice can be made (and documented) amongst the alternatives.

The ease with which this translation can be defined will of course depend on the clarity with which scheme X represents the features it encodes.

Translating from the TEI into scheme X follows the same pattern, except that if a TEI feature has no equivalent in X, and X cannot be extended, information must be lost in translation.

The rules defining conformance to these Guidelines are given in some detail in chapter *24.4. Conformance*. The basic principles informing those rules may be summarized as follows:

1. The TEI *abstract model* (that is, the set of categorical distinctions which it defines in the prose of the Guidelines) must be respected. The correspondence between a tag X and the semantic function assigned to it by these Guidelines may not be changed; such changes are known as *tag abuse* and strongly discouraged.
2. A TEI document must be expressed as a valid XML-conformant document which uses the TEI namespace appropriately. If, for example, the document encodes features not provided by these Guidelines, such extensions should not be associated with the TEI namespace.
3. It must be possible to validate a TEI document against a schema derived from these Guidelines, possibly with extensions provided in the recommended manner.

### Use for Local Processing

Machine-readable text can be manipulated in many ways; some users:

- edit texts (e.g. word processors, syntax-directed editors)
- edit, display, and link texts in hypertext systems
- format and print texts using desktop publishing systems, or batch-oriented formatting programs
- load texts into free-text retrieval databases or conventional databases
- unload texts from databases as search results or for export to other software
- search texts for words or phrases
- perform content analysis on texts
- collate texts for critical editions
- scan texts for automatic indexing or similar purposes
- parse texts linguistically
- analyze texts stylistically
- scan verse texts metrically
- link text and images

These applications cover a wide range of likely uses but are by no means exhaustive. The aim has been to make the TEI Guidelines useful for encoding the same texts for different purposes. We have avoided anything which would restrict the use of the text for other applications. We have also tried not to omit anything essential to any single application.

Because the TEI format is expressed using XML, almost any modern text processing system is able to process it, and new TEI-aware software systems are able to build on a solid base of existing software libraries.

## iv.2 Historical Background

The Text Encoding Initiative grew out of a planning conference sponsored by the Association for Computers and the Humanities (ACH) and funded by the U.S. National Endowment for the Humanities (NEH), which was held at Vassar College in November 1987. At this conference some thirty representatives of text archives, scholarly societies, and research projects met to discuss the feasibility of a standard encoding scheme and to make recommendations for its scope, structure, content, and drafting. During the conference, the Association for Computational Linguistics and the Association for Literary and Linguistic Computing agreed to join ACH as sponsors of a project to develop these Guidelines. The outcome of the conference was a set of principles (the 'Poughkeepsie Principles', ?), which determined the further course of the project.

The Text Encoding Initiative project began in June 1988 with funding from the NEH, soon followed by further funding from the Commission of the European Communities, the Andrew W. Mellon Foundation, and the Social Science and Humanities Research Council of Canada. Four working committees, composed of distinguished scholars and researchers from both Europe and North America, were named to deal with problems of text documentation, text representation, text analysis and interpretation, and metalanguage and syntax issues. Each committee was charged with the task of identifying 'significant particularities' in a range of texts, and two editors appointed to harmonize the resulting recommendations.

A first draft version (P1, with the 'P' here and subsequently standing for 'Proposal') of the Guidelines was distributed in July 1990 under the title *Guidelines for the Encoding and Interchange of Machine-Readable Texts*. Extensive public comment and further work on areas not covered in this version resulted in the drafting of a revised version, TEI P2, distribution of which began in April 1992. This version included substantial amounts of new material, resulting from work carried out by several specialist working groups, set up in 1990 and 1991 to propose extensions and revisions to the text of P1. The overall organization, both of the draft itself and of the scheme it describes, was entirely revised and reorganized in response to public comment on the first draft.

In June 1993 an Advisory Board met to review the current state of the TEI Guidelines, and recommended the formal publication of the work done to that time. That version of the TEI Guidelines, TEI P3, consolidated the work published as parts of TEI P2, along with some additional new material and was finally published in May of 1994 without the label *draft*, thus marking the conclusion of the initial development work.

In February of 1998 the World Wide Web Consortium issued a final Recommendation for the Extensible Markup Language, XML.[2] Following the rapid take-up of this new standard metalanguage, it became evident that the TEI Guidelines (which had been published originally as an SGML application) needed to be re-expressed in this new formalism if they were to survive. The TEI editors, with abundant assistance from others who had developed and used TEI, developed an update plan, and made tentative decisions on relevant syntactic issues.

In January of 1999, the University of Virginia and the University of Bergen formally proposed the creation of an international membership organization, to be known as the TEI Consortium, which would maintain, develop, and promote the TEI. Shortly thereafter, two further institutions with longstanding ties to the TEI (Brown University and Oxford University) joined them in formulating an Agreement to Establish a Consortium for the Maintenance of the Text Encoding Initiative (An Agreement to Establish a Consortium for the Maintenance of the Text Encoding Initiative (March 1999)),
on which basis the TEI Consortium was eventually established and incorporated as a not-for-profit legal entity at the end of the year 2000. The first members of the new TEI Board took office during January of 2001.

The TEI Consortium was established in order to maintain a permanent home for the TEI as a democratically constituted, academically and economically independent, self-sustaining, non-profit organization. In addition, the TEI Consortium was intended to foster a broad-based user community with sustained involvement in the future development and widespread use of the TEI Guidelines (Burnard (2000)).

To oversee and manage the revision process in collaboration with the TEI Editors, the TEI Board formed a Technical Council, with a membership elected from the TEI user community. The Council met for the first time in January 2002 at King's College London. Its first task was to oversee production of an XML version of the TEI Guidelines, updating P3 to enable users to work with the emerging XML toolset. This, the P4 version of the Guidelines, was published in June 2002. It was essentially an XML version of P3, making no substantive changes to the constraints expressed in the schemas apart from those necessitated by the shift to XML, and changing only

[2] XML was originally developed as a way of publishing on the World Wide Web richly encoded documents such as those for which the TEI was designed. Several TEI participants contributed heavily to the development of XML, most notably XML's senior co-editor C. M. Sperberg-McQueen, who served as the North American editor for the TEI Guidelines from their inception until 1999.

corrigible errors identified in the prose of the P3 Guidelines. However, given that P3 had by this time been in steady use since 1994, it was clear that a substantial revision of its content was necessary, and work began immediately on the P5 version of the Guidelines. This was planned as a thorough overhaul, involving a public call for features and new development in a number of important areas not previously addressed including character encoding, graphics, manuscript description, biographical and geographical data, and the encoding language in which the TEI Guidelines themselves are written.

The members of the TEI Council and its associated workgroups are listed in *iii Preface and Acknowledgments*. In preparing this edition, they have been attentive to the requirements and practice of the widest possible range of TEI users, who are now to be found in many different research communities across the world, and have been largely instrumental in transforming the TEI from a grant-supported international research project into a self-sustaining community-based effort. One effect of the incorporation of the TEI has been the legal requirement to hold an annual meeting of the Consortium members; these meetings have emerged as an invaluable opportunity to sustain and reinforce that sense of community.

The present work is therefore the result of a sustained period of consultation, drafting, and revision, with input from many different experts. Whatever merits it may have are to be attributed to them; the Editors accept responsibility only for the errors remaining.

## iv.3 Future Developments and Version Numbers

The encoding recommended by this document may be used without fear that future versions of the TEI scheme will be inconsistent with it in fundamental ways. The TEI will be sensitive, in revising these Guidelines, to the possible problems which revision might pose for those who are already using this version of these Guidelines.

With TEI P5, a version numbering system is introduced following the pattern specified by the Unicode Consortium: the first digit identifies a major version number, the second digit a minor version number, and the third digit a sub-minor version number. The TEI undertakes that no change will be made to the formal expression of these Guidelines (that is, a TEI schema, as defined in *24.4. Conformance*) such that documents conformant to a given major numbered release cease to be compatible with a subsequent release of the same major number. Moreover, as far as possible, new minor releases will be made only for the purpose of adding new compatible features, or of correcting errors in existing features.

The Guidelines are currently maintained as an open source project on the Github site `https://github.com/TEIC/TEI`, from which released and development versions may be freely downloaded. See Previous Releases of P5 for information on how to find specific versions of TEI releases (Guidelines, schemas etc.). Notice of errors detected and enhancements requested may be submitted at `https://github.com/TEIC/TEI/issues`.

# v A Gentle Introduction to XML

The encoding scheme defined by these Guidelines is formulated as an application of the Extensible Markup Language (XML) (?). XML is widely used for the definition of device-independent, system-independent methods of storing and processing texts in electronic form. It is now also the interchange and communication format used by many applications on the World Wide Web. In the present chapter we informally introduce some of its basic concepts and attempt to explain to the reader encountering them for the first time how and why they are used in the TEI scheme. More detailed technical accounts of TEI practice in this respect are provided in chapters *24. Using the TEI*, *1. The TEI Infrastructure*, and *23. Documentation Elements* of these Guidelines.

Strictly speaking, XML is a *metalanguage*, that is, a language used to describe other languages, in this case, *markup* languages. Historically, the word *markup* has been used to describe annotation or other marks within a text intended to instruct a compositor or typist how a particular passage should be printed or laid out. Examples include wavy underlining to indicate boldface, special symbols for passages to be omitted or printed in a particular font, and so forth. As the formatting and printing of texts was automated, the term was extended to cover all sorts of special codes inserted into electronic texts to govern formatting, printing, or other processing.

Generalizing from that sense, we define *markup*, or (synonymously) *encoding*, as any means of making explicit an interpretation of a text. Of course, all printed texts are implicitly encoded (or marked up) in this sense: punctuation marks, capitalization, disposition of letters around the page, even the spaces between words all might be regarded as a kind of markup, the purpose of which is to help the human reader determine where one word ends and another begins, or how to identify gross structural features such as headings or simple syntactic units such as dependent clauses or sentences. Encoding a text for computer processing is, in principle, like transcribing a manuscript from *scriptio continua*[3]; it is a process of making explicit what is conjectural or implicit, a process of directing the user as to how the content of the text should be (or has been) interpreted.

By *markup language* we mean a set of markup conventions used together for encoding texts. A markup language must specify how markup is to be distinguished from text, what markup is allowed, what markup is required, and what the markup means. XML provides the means for doing the first three; documentation such as these Guidelines is required for the last.

The present chapter attempts to give an informal introduction to those parts of XML of which a proper understanding is necessary to make best use of these Guidelines. The interested reader should also consult one or more of the many excellent introductory textbooks and web sites now available on the subject.[4]

## v.1 What's Special about XML?

XML has three highly distinctive advantages:

1. it places emphasis on descriptive rather than procedural markup;
2. it distinguishes the concepts of syntactic correctness and of *validity* with respect to a *document type definition*;
3. it is independent of any one hardware or software system.

These three aspects are discussed briefly below, and then in more depth in the remainder of this chapter.

XML is frequently compared with HTML, the language in which web pages have generally been written, which shares some of the above characteristics. Compared with HTML, however, XML has some other important features:

- XML is *extensible*: it does not consist of a fixed set of tags;
- XML documents must be *well-formed* according to a defined syntax;
- an XML document can be formally *validated* against a set of schema rules for consistent application;
- XML is more interested in the meaning of data than in its presentation.

[3] In the 'continuous writing' characteristic of manuscripts from the early classical period, words are written continuously with no intervening spaces or punctuation.

[4] New textbooks and websites about XML appear at regular intervals and to select any one of them would be invidious. some recommended online courses include `http://www.w3schools.com/xml/default.asp` and `https://www.ibm.com/developerworks/xml/tutorials/xmlintro/xmlintro.html`.

### v.1.1 Descriptive Markup

In a descriptive markup system, the markup codes used do little more than categorize parts of a document. Markup codes such as <para> or `\end{list}` simply identify a portion of a document and assert of it that 'the following item is a paragraph', or 'this is the end of the most recently begun list', etc. By contrast, a procedural markup system defines what processing is to be carried out at particular points in a document: 'call procedure PARA with parameters 42, b, and x here' or 'move the left margin 2 quads left, move the right margin 2 quads right, skip down one line, and go to the new left margin,' etc. In XML, the instructions needed to process a document for some particular purpose (for example, to format it) are sharply distinguished from the markup used to describe it.

Usually, the markup or other information needed to process a document will be maintained separately from the document itself, typically in a distinct document called a *stylesheet*, though it may do much more than simply define the rendition or visual appearance of a document.[5]

When descriptive markup is used, the same document can readily be processed in many different ways, using only those parts of it which are considered relevant. For example, a content analysis program might disregard entirely the footnotes embedded in an annotated text, while a formatting program might extract and collect them all together for printing at the end of each chapter. Different kinds of processing can be carried out with the same part of a file. For example, one program might extract names of persons and places from a document to create an index or database, while another, operating on the same text, but using a different stylesheet, might print names of persons and places in a distinctive typeface.

### v.1.2 Types of Document

A second key aspect of XML is its notion of a *document type*: documents are regarded as having types, just as other objects processed by computers do. The type of a document is formally defined by its constituent parts and their structure. The definition of a 'report', for example, might be that it consisted of a 'title' and possibly an 'author', followed by an 'abstract' and a sequence of one or more 'paragraphs'. Anything lacking a title, according to this formal definition, would not formally be a report, and neither would a sequence of paragraphs followed by an abstract, whatever other report-like characteristics these might have for the human reader.

If documents are of known types, a special-purpose program (called a *parser*), once provided with an unambiguous definition of a document type, can check that any document claiming to be of that type does in fact conform to the specification. A parser can check that all elements specified for a particular document type are present and no others, that they are combined in appropriate ways, correctly ordered, and so forth. More significantly, different documents of the same type can be processed in a uniform way. Programs can be written which take advantage of the knowledge encapsulated in the document type information, and which can thus behave in a more 'intelligent' fashion.

### v.1.3 Data Independence

A basic design goal of XML is to ensure that documents encoded according to its provisions can move from one hardware and software environment to another without loss of information. The two features discussed so far both address this requirement at an abstract level; the third feature addresses it at the level of the strings of data characters that make up a document. All XML documents, whatever languages or writing systems they employ, use the same underlying character encoding (that is, the same method of representing as binary data those graphic forms making up a particular writing system).[6] This encoding is defined by an international standard,[7] which is implemented by a universal character set maintained by an industry group called the Unicode Consortium, and known as Unicode.[8] Unicode provides a standardized way of representing any of the many thousands of discrete symbols making up the world's writing systems, past and present.

Most modern computing systems now support Unicode directly; for those which do not, XML provides a mechanism for the indirect representation of single characters by means of their character number, known as *character references*; see further *v.7.1 Character References*.

[5] We do not here discuss in any detail the ways that a stylesheet can be used or defined, nor do we discuss the popular W3C Stylesheet Languages XSLT and CSS. See further ?, ?, and ?.

[6] See *Extensible Markup Language (XML) 1.0*, available from `https://www.w3.org/TR/REC-xml/`, Section 2.2 Characters.

[7] ISO/IEC 10646-1993 *Information Technology — Universal Multiple-Octet Coded Character Set* (UCS)

[8] See `https://www.unicode.org/`

# v.2 Textual Structures

A text is not an undifferentiated sequence of words, much less of bytes. For different purposes, it may be divided into many different units, of different types or sizes. A prose text such as this one might be divided into sections, chapters, paragraphs, and sentences. A verse text might be divided into cantos, stanzas, and lines. Once printed, sequences of prose and verse might be divided into volumes, gatherings, and pages.

Structural units of this kind are most often used to identify specific locations or refer to points within a text ('the third sentence of the second paragraph in chapter ten'; 'canto 10, line 1234'; 'page 412', etc.) but they may also be used to subdivide a text into meaningful fragments for analytic purposes ('is the average sentence length of section 2 different from that of section 5?' 'how many paragraphs separate each occurrence of the word *nature*? how many pages?'). Other structural units are more clearly analytic, in that they characterize a section of a text. A dramatic text might regard each speech by a different character as a unit of one kind, and stage directions or pieces of action as units of another kind. Such an analysis is less useful for locating parts of the text ('the 93rd speech by Horatio in Act 2') than for facilitating comparisons between the words used by one character and those of another, or those used by the same character at different points of the play.

In a prose text one might similarly wish to regard as units of different types passages in direct or indirect speech, passages employing different stylistic registers (narrative, polemic, commentary, argument, etc.), passages of different authorship and so forth. And for certain types of analysis (most notably textual criticism) the physical appearance of one particular printed or manuscript source may be of importance: paradoxically, one may wish to use descriptive markup to describe presentational features such as typeface, original topographic lineation, use of whitespace, and so forth.

These textual structures overlap with one another in complex and unpredictable ways. Particularly when dealing with texts as instantiated by paper technology, the reader needs to be aware of both the physical organization of the book and the logical structure of the work it contains. Many great works (Sterne's *Tristram Shandy* for example) cannot be fully appreciated without an awareness of the interplay between narrative units (such as chapters or paragraphs) and presentational ones (such as page divisions). For many types of research, the interplay among different levels of analysis is crucial: the extent to which syntactic structure and narrative structure mesh, or fail to mesh, for example, or the extent to which phonological structures reflect morphology.

# v.3 XML Structures

This section describes the simple and consistent mechanism for the markup or identification of textual structure provided by XML. It also describes the methods XML provides for the expression of rules defining how units of textual structure can meaningfully be combined in a text.

## v.3.1 Elements

The technical term used in XML for a textual unit, viewed as a structural component, is *element*. Different types of elements are given different names, but XML provides no way of expressing the meaning of a particular type of element, other than its relationship to other element types. That is, all one can say about an element called (say) `<blort>` is that instances of it may (or may not) occur within elements of type `<farble>`, and that it may (or may not) be decomposed into elements of type `<blortette>`. It should be stressed that XML is entirely unconcerned with the *semantics* of textual elements, because these are considered to be application dependent. It is up to the creators of XML vocabularies (such as these Guidelines) to choose intelligible element names and to define their intended use in text markup. That is the chief purpose of documents such as the TEI Guidelines. From the need to choose element names indicative of function comes the technical term for the name of an element type, which is *generic identifier*, or GI.

Within a marked-up text (a *document instance*), each element must be explicitly marked or tagged in some way. This is done by inserting a tag at the beginning of the element (a *start-tag*) and another at its end (an *end-tag*). The start- and end-tag pair are used to bracket off element occurrences within the running text, in rather the same way as different types of parentheses or quotation marks are used in conventional punctuation. For example, a quotation element in a text might be tagged as follows:

```
... Rosalind's
remarks <quote>This is the silliest stuff that ere I heard
of!</quote> clearly indicate ...
```

As this example shows, a start-tag takes the form <quote>, where the opening angle bracket indicates the start of the start-tag, 'quote' is the generic identifier of the element that is being delimited, and the closing angle bracket indicates the end of the start-tag. An end-tag takes an identical form, except that the opening angle bracket is followed by a solidus (slash) character, so that the corresponding end-tag is </quote>.[9] The material between the start-tag and the end-tag (the string of words 'This is the silliest stuff that ere I heard of' in the example above) is known as the *content* of the element. Sometimes there may be nothing between the start and the end-tag; in this case the two may optionally be merged together into a single composite tag with the solidus at the end, like this: <quote/>.

### v.3.2 Content Models: an Example

An element may be *empty*, that is, it may have no content at all, or it may contain just a sequence of characters with no other elements. Often, however, elements of one type will be *embedded* (contained entirely) within elements of a different type.

To illustrate this, we will consider a very simple structural model. Let us assume that we wish to identify within an anthology only poems, their headings, and the stanzas and lines of which they are composed. In XML terms, our *document type* is the *anthology*, and it consists of a series of *poems*. Each poem has embedded within it one element, a *heading*, and several occurrences of another, a *stanza*, each stanza having embedded within it a number of *line* elements. Fully marked up, a text conforming to this model might appear as follows:

```
<anthology>
  <poem>
    <heading>The SICK ROSE</heading>
    <stanza>
      <line>O Rose thou art sick.</line>
      <line>The invisible worm,</line>
      <line>That flies in the night</line>
      <line>In the howling storm:</line>
    </stanza>
    <stanza>
      <line>Has found out thy bed</line>
      <line>Of crimson joy:</line>
      <line>And his dark secret love</line>
      <line>Does thy life destroy.</line>
    </stanza>
  </poem>
<!-- more poems go here -->
</anthology>
```

It should be stressed that this example does *not* use the names proposed for corresponding elements elsewhere in these Guidelines: the above is thus *not* a valid TEI document.[10] It will, however, serve as an introduction to the basic notions of XML. Whitespace and line divisions have been added to the example for the sake of visual clarity only; they have no particular significance in the XML encoding itself. Also, the line

```
  <!-- more poems go here -->
```

is an XML *comment* and is not treated as part of the text.

As it stands, the above example is what is known as a *well-formed* XML document because it obeys the following simple rules:

1. there is a single element enclosing the whole document: this is known as the *root element* (`<anthology>` in our case);

[9] Because the opening angle bracket has this special function in an XML document, special steps must be taken to use that character for other purposes (for example, as the mathematical less-than operator); see further section *v.7.1 Character References*.

[10] The element names here have been chosen for clarity of exposition; there is, however, a TEI element corresponding to each.

2. each element is completely contained by the root element, or by an element that is so contained; elements do not partially overlap one another;
3. a tag explicitly marks the start and end of each element.

A well-formed XML document can be processed in a number of useful ways. A simple indexing program could extract only the relevant text elements in order to make a list of headings, first lines, or words used in the poem text; a simple formatting program could insert blank lines between stanzas, perhaps indenting the first line of each, or inserting a stanza number. Different parts of each poem could be typeset in different ways. A more ambitious analytic program could relate the use of punctuation marks to stanzaic and metrical divisions.[11] Scholars wishing to see the implications of changing the stanza or line divisions chosen by the editor of this poem can do so simply by altering the position of the tags. And of course, the text as presented above can be transported from one computer to another and processed by any program (or person) capable of making sense of the tags embedded within it with no need for the sort of transformations and translations needed for files which have been saved in one or other of the proprietary formats preferred by most word-processing programs.

As we noted above, one of the attractions of XML is that it enables us to apply our own names for the elements rather than requiring us always to use names predefined by other agencies. Clearly, however, if we wish to exchange our poems with others, or to include poems others have marked up in our anthology, we will need to know a bit more about the names used for the tags. The means that XML provides for this is called a *namespace*. In our simple example, the tags just contain a simple name. As we shall see, it is also possible to use tags that include a *qualified name*, that is, a name with an optional prefix identifying the set of names to which it belongs. For example, we have defined an element `<line>` for the purpose of marking lines of verse. Another person might, however, define an element called `<line>` for the purpose of marking typographic lines, or drawn lines. Because of these different meanings, if we wish to share data it will be necessary to distinguish the two 'line' components in our marked-up texts. This is achieved by including a *namespace prefix* within the markup, for example like this:

```
  <my:line>This is one of my lines</my:line>
<!-- ... -->
<yr:line>This is one of your lines</yr:line>
```

This feature is particularly important if we have different definitions of what a 'line' is, of course, but there are many occasions when it is useful to distinguish groups of tags belonging to different 'markup vocabularies'; we discuss this further below (*v.7.2 Namespaces*). One particularly useful namespace prefix is predefined for XML: it is `xml` and we will see examples of its use below.

Namespaces allow us to represent the fact that a name belongs to a group of names, but don't allow us to do much more by way of checking the integrity or accuracy of our tagging. Simple well-formedness alone is not enough for the full range of what might be useful in marking up a document. It might well be useful if, in the process of preparing our digital anthology, a computer system could check some basic rules about how stanzas, lines, and headings can sensibly co-occur in a document. It would be even more useful if the system could check that stanzas are always tagged <stanza> and not occasionally <canto> or <Stanza>. An XML document in which such rules have been checked is technically known as a *valid* document, and the ability to perform such validation is one of the key advantages of using XML. To carry this out, some way of formally stating the criteria for successful validation is necessary: in XML this formal statement is provided by an additional document known as a *schema*.[12]

## v.4 Validating a Document's Structure

The design of a schema may be as lax or as restrictive as the occasion warrants. A balance must be struck between the convenience of following simple rules and the complexity of handling real texts. This is particularly the case when the rules being defined relate to texts that already exist: the designer may have only the haziest of notions as to an ancient text's original purpose or meaning and hence find it very difficult to specify consistent rules about its

[11] Note that this simple example has not addressed the problem of marking elements such as sentences explicitly; the implications of this are discussed in section *v.5 Complicating the Issue*.

[12] The older terms *Document Type Declaration* and *Document Type Definition*, both abbreviated as DTD, may also be encountered. Throughout these Guidelines we use the term *schema* for any kind of formal document grammar.

structure. On the other hand, where a new text is being prepared to an exact specification, for entry into a textual database of some kind for example, the more precisely stated the rules, the better they can be enforced. Even in marking up an existing text, a restrictive set of schema rules may be beneficial, especially when applied to test a particular view or hypothesis about the text. A schema designed for use by a small project or team is likely to take a different position on such issues than one intended for use by a large and possibly fragmented community. It is important to remember that every schema results from an interpretation of a text. There is no single schema encompassing the absolute truth about any text, although it may be convenient to privilege some schemas above others for particular types of analysis.

XML is widely used in environments where uniformity of document structure is a major desideratum. In the production of technical documentation, for example, it is of major importance that sections and subsections should be properly nested, that cross-references should be properly resolved and so forth. In such situations, documents are seen as raw material to match against predefined sets of rules. As discussed above, however, the use of simple rules can also greatly simplify the task of tagging accurately elements of less rigidly constrained texts. By making these schema rules explicit, scholars reduce their own burdens with consistently marking up and verifying the electronic text. By defining and sharing their schema rules, scholars openly express a project-specific interpretation of the structure and significant particularities of the text being encoded.

Schema validation for XML is usually written in the RELAX NG language (`http://relaxng.org/`) originally developed within the OASIS Technical Committee and now an ISO standard[13], though other older methods include the Document Type Definition (DTD) language which XML inherited from SGML and the XML Schema language (`https://www.w3.org/XML/Schema`) defined by the W3C.[14] In this chapter, and throughout these Guidelines, we give examples using the 'compact syntax' of RELAX NG for ease of reading. The specifications for the TEI Guidelines are first expressed in the TEI language itself and a RELAX NG schema is generated from them for processing convenience. Details about schema customization using the TEI ODD language are addressed in *23. Documentation Elements*, *24.3. Customization* and *24.5. Implementation of an ODD System*.

## v.4.1 An Example Schema

For the purposes of illustrating how a schema works to restrict how XML may be written, we use the RELAX NG compact syntax in what follows. The following schema can be used to validate our example poem:

```
  anthology_ p = element anthology { poem_ p+ }
poem_ p = element poem { heading_ p?, stanza_ p+ }
stanza_ p = element stanza {line_ p+}
heading_ p = element heading { text }
line_ p = element line { text }
start = anthology_ p
```

Note that this is not the only way in which a RELAX NG schema might be written; we have adopted this idiom, however, because it matches that used throughout the rest of these Guidelines.

A RELAX NG schema expresses rules about the possible structure of a document in terms of *patterns*; that is, it defines a number of named patterns, each of which acts as a kind of template against which an input document can be matched. The meaning of a pattern is expressed in a schema either by reference to other patterns, or to a small number of fundamental built-in concepts, as we shall see. In the example above, the word to the left of the equals sign is the pattern's name, and the material following it declares a meaning for the pattern. Patterns may also be of particular types; the ones that interest us here are called *element patterns* and *attribute patterns*. In this example we see definitions for five element patterns. Note that we have used similar names for the pattern and the element which the pattern describes: so, for example, the line `anthology_p = element anthology {poem_p+}` defines an element pattern called `anthology_p`, the value of which defines an element called `anthology`. These naming conventions are arbitrary; we could use the same name for the pattern as for the element, but we want to make clear that the two are syntactically quite distinct. The name, or *generic identifier*, of the element follows the

[13] ISO/IEC FDIS 19757-2 Document Schema Definition Language (DSDL) — Part 2: Regular-grammar-based validation — RELAX NG

[14] Schema validation languages co-evolved with early markup language specifications, as summarized in Rick Jelliffe's Family Tree of Schema Languages for Markup Languages.

word 'element', and the *content model* for the element is given within the curly braces following that. Each of these parts is discussed further below.

The last line of the schema above tells a RELAX NG validator which element (or elements) in a document can be used as the root element: in our case only `<anthology>`. This enables the validator to detect whether a particular document is well-formed but incomplete; it also simplifies the processing task by providing an 'entry point'.

### Generic Identifier

Following the word 'element' each pattern declaration gives the generic identifier (often abbreviated to GI) of the element being defined, for example *poem*, *heading*, etc. A GI may contain letters, digits, hyphens, underscore characters, or full stops, but must begin with a letter and may not contain a space.[15] Uppercase and lowercase letters are quite distinct: an element with the GI `<foo>` is *not* the same as an element with the GI `<Foo>`; the root element of a TEI-conformant document is <TEI>, *not* `<tei>`.

### Content Model

The second part of each declaration, enclosed in curly braces, is called the *content model* of the element being defined, because it specifies what may legitimately be contained within it. In RELAX NG, the content model is defined in terms of other patterns, either by embedding them, or (as in our examples above) by naming or referring to them. The RELAX NG compact syntax also uses a small number of reserved words to identify other possible contents for an element, of which by far the most commonly encountered is `text`, as in this example: it means that the element being defined may contain any valid character data, but no elements. If an XML document is thought of as a structure like a family tree, with a single ancestor at the top (in our case, this would be <anthology>), then almost always, following the branches of the tree downwards (for example, from `<anthology>` to `<poem>` to `<stanza>` to `<line>` and `<heading>`) will lead eventually to `text`. In our example, `<heading>` and `<line>` are so defined, since their content models say `text` only and name no embedded elements.

### Occurrence Indicators

The declaration for `<stanza>` in the example above states that a stanza consists of one or more lines. It uses an *occurrence indicator* (the plus sign) to indicate how many times something matching the pattern `line_p` may be repeated. There are three occurrence indicators: the plus sign, the question mark, and the asterisk or star. The plus sign means that the pattern can match one or more times; the question mark means that it may match at most once but is not mandatory; the star means that the pattern concerned is not mandatory, but may match more than once. Thus, if the content model for `<stanza>` were `{line_p*}`, stanzas with no lines would be possible as well as those with more than one line. If it were `{line_p?}`, again empty stanzas would be permitted, but no stanza could have more than a single line. The declaration for `<poem>` in the example above thus states that a `<poem>` cannot have more than one heading, but may have none, and that it must have at least one `<stanza>` and may have several.

### Connectors

The content model `{heading_p?, stanza_p+}` contains more than one component, and thus needs additionally to specify the order in which these patterns (`<heading_p>` and `<stanza_p>`) may appear. This ordering is determined by the *connector* (the comma) used between its components. The comma connector indicates that the patterns concerned must appear in the sequence given. Another commonly encountered connector is the vertical bar, representing alternation. If the comma in this example were replaced by a vertical bar, then a `<poem>` would consist of either a heading or just stanzas—but not both!

### Groups

In our example so far, the components of each content model have been either single patterns or `text`. We often need to define content models in more complicated ways, in which the components are lists of patterns, combined by connectors. Such lists may also be modified by occurrence indicators and themselves combined by connectors. To demonstrate these facilities, let us expand our example so that it may include non-stanzaic types of verse. For the sake of demonstration, we will categorize poems as one of the following: *stanzaic*, *couplets*, or *blank* (or *stichic*). A blank-verse poem consists simply of lines (we ignore the possibility of verse paragraphs for the moment), so no additional

[15] In XML, a single colon may also appear in a GI, where it has a special significance related to the use of *namespaces*, as further discussed in section *v.7.2 Namespaces*. The characters defined by Unicode as *combining characters* and as *extenders* are also permitted, as are logograms such as Chinese characters.

elements need be defined for it. We could define a couplet as a <firstLine> followed by a <secondLine>, which distinction might be useful in a study of rhyme schemes.[16]

```
  couplet_ p = element couplet {firstLine_ p, secondLine_ p}
```

The patterns firstLine_p and secondLine_p define elements <firstLine> and <secondLine>; these will have exactly the same content model as the existing <line> element. We will therefore add the following two lines to our example schema:

```
  firstLine_ p = element firstLine {text} secondLine_ p = element secondLine {text}
```

Next, we can change the declaration for the <poem> element to include all three possibilities:

```
  poem_ p = element poem
{ heading_ p?, (stanza_ p+ | couplet_ p+ | line_ p+) }
```

That is, a poem consists of an optional heading, followed by one or several stanzas, or one or several couplets, or one or several lines. Note the difference between this declaration and the following:

```
  poem_ p = element poem
{heading_ p?, (stanza_ p | couplet_ p | line_ p)+ }
```

The second version, by applying the occurrence indicator to the group rather than to each element within it, would allow a single poem to contain a mixture of stanzas, couplets, and lines.

A group of this kind can contain text as well as named elements: this combination, known as *mixed content*, allows for elements in which the sub-components appear with intervening stretches of character data. For example, if we wished to mark place names wherever they appear inside our verse lines, then, assuming we have also added a pattern for the <name> element, we could change the definition for <line> to

```
  line_ p = element
line { (text | name_ p )* }
```

Some XML schema languages place no constraints on the way that mixed content models may be defined, but in the XML DTD language, when text appears with other elements in a content model, it must always appear as the first option in an alternation; it may appear once only, and in the outermost model group; and if the group containing it is repeated, the star operator must be used. Although these constraints are not strictly necessary in RELAX NG schemas, all TEI content models currently obey them.

Quite complex models can be built in this way, to match the structural complexity of many types of text. For example, consider the case of stanzaic verse in which a refrain or chorus appears. Like a stanza, a refrain consists of repetitions of the line element. A refrain can appear at the start of a poem only, or as an optional addition following each stanza. This could be expressed by a pattern such as the following:

[16] This example is probably not a good practice for most XML projects, since XPath provides ways of distinguishing elements in an XML structure by their position, or the order in which they appear in relation to one another, without the need to give them distinct names.

```
  refrain_ p = element refrain {line_ p+}
poem_ p = element poem   {heading_ p?, ( line_ p+ | (refrain_ p?, (stanza_ p,
refrain_ p?)+ )) }
```

That is, a poem consists of an optional heading, followed by either a sequence of lines or an unnamed group, which starts with an optional refrain and is followed by one or more occurrences of another group, each member of which is composed of a stanza followed by an optional refrain. A sequence such as *refrain - stanza - stanza - refrain* follows this pattern, as does the sequence *stanza - refrain - stanza - refrain.* The sequence *refrain - refrain - stanza - stanza* does not, however, and neither does the sequence *stanza - refrain - refrain - stanza.* Among other conditions made explicit by this content model are the requirements that at least one stanza must appear in a poem, if it is not composed simply of lines, and that if there is both a heading and a stanza they must appear in that order.

Note that the apparent complexity of this model derives from the constraints expressed informally above. A simpler model, such as

```
  poem_ p =
element poem {heading_ p?, (line_ p | refrain_ p | stanza_ p)+ }
```

would not enforce any of them, and would therefore permit such anomalies as a poem consisting only of refrains, or an arbitrary mixture of lines and refrains.

It is beyond the scope of this "Gentle Introduction to XML" to provide a complete orientation to schema writing with Relax NG, but interested readers may wish to consult more thorough tutorials on the subject.[17] The examples and explanation provided here may be helpful to consult when reading the schema declarations posted for groupings of TEI elements that share the same content model, such as `macro.phraseSeq`, whose declaration features an example of mixed text and element content. Due to the complexity of the TEI schema as expressed in Relax NG, it is best to customize its content model in TEI itself by writing an ODD, as discussed in *24.5. Implementation of an ODD System.*

## v.5 Complicating the Issue

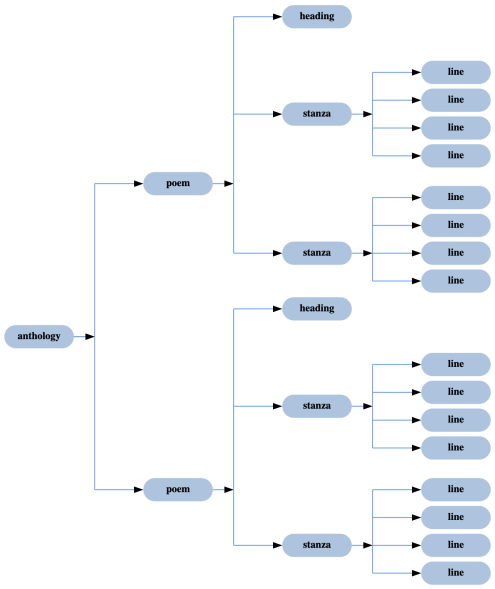
In the simple cases described so far, we have assumed that one can identify the immediate constituents of every element in a textual structure. A poem consists of stanzas, and an anthology consists of poems. Stanzas do not float around unattached to poems or combined into some other unrelated element; a poem cannot contain an anthology. All the elements of a given document type may be arranged into a hierarchic structure like a family tree, with a single ancestor at one end and many children (mostly the elements containing simple text) at the other. For example, we could represent an anthology containing two poems, the first of which contains two four-line stanzas and the second a single stanza, by a tree structure like the following figure:

This graphic represents the hierarchical structure of an XML document, resembling a family tree. Most XML processing systems now use a standardized way of accessing parts of an XML document called *XPath.*[18] XPath gives us a non-graphical way of referring to any part of an XML document: for example, we might refer to the last line of Blake's poem as `/anthology/poem[1]/stanza[2]/line[4]`. The square brackets here indicate a numerical selection: we are talking about the fourth line in the second stanza of the first poem in the anthology. If we left out all the square-bracketted selections, the corresponding XPath expression would refer to all lines contained by stanzas contained by poems contained by anthologies. An XPath expression can refer to any collection of elements: for example, the expression `/anthology/poem` refers to all poems in an anthology and the expression `/anthology/poem/heading` refers to all their headings.

The forward slash ('/', U+002F SOLIDUS) within an XPath expression behaves in much the same way as a forward slash or backslash does in a filename specification. To use a family tree analogy, a single slash indicates that the item to the immediate left is a parent of the item(s) to the right of it. For example, in the XPath expression

[17] For a complete tutorial introduction to RELAX NG, see ?.

[18] The official specification is at ?; many introductory tutorials are available in the XML references cited above and elsewhere on the Web: good beginners' tutorials include `http://dh.obdurodon.org/introduction-xpath.xhtml`, `http://www.w3schools.com/xml/xpath_intro.asp` and `http://www.zvon.org/xxl/XPathTutorial/`, the latter being available in several languages.




`/anthology/poem`, the single slash between anthology and poem indicates that anthology is a parent of the poem children elements. (The first forward slash in the XPath expression indicates the document node.) In XPath, it is also possible to refer to children, grandchildren, and other descendants of the family tree using two forward slashes together. For example, the XPath expression `/anthology/poem//line` will refer to all of the lines of all of the stanzas of all the poems, without having to represent the stanza element in the XPath.

Clearly, there are many such trees that might be drawn to describe the structure of this or other anthologies. Some of them might be representable as further subdivisions of this tree: for example, we might subdivide the lines into individual words, since in our simple example no word crosses a line boundary. Surprisingly perhaps, this grossly simplified view of what text is (memorably termed an *ordered hierarchy of content objects* (OHCO) view of text by Renear *et al.*[19]) turns out to be very effective for a large number of purposes. It is not, however, adequate for the full complexity of real textual structures, for which more complex mechanisms need to be employed. There are many other trees that might be drawn which do *not* fit within the anthology model which we have presented so far. We might, for example, be interested in syntactic structures or other linguistic constructs, which rarely respect the formal boundaries of verse. Or, to take a simpler example, we might want to represent the pagination of different editions of the same text.

In the OHCO model of text, representation of cases where different elements overlap so that several different trees may be identified in the same document is generally problematic. All the elements marked up in a document, no matter what namespace they belong to, must fit within a single hierarchy. To represent overlapping structures, therefore, a single hierarchy must be chosen, and the points at which other hierarchies intersect with it marked. For example, we might choose the verse structure as our primary hierarchy, and then mark the pagination by means of empty elements inserted at the boundary points between one page and the next. Or we could represent alternative hierarchies by means of the pointing and linking mechanisms described in chapter *17. Linking, Segmentation, and Alignment* of these Guidelines. These mechanisms all depend on the use of *attributes*, which may be used both to identify particular elements within a document and to point to, link, or align them into arbitrary structures.

## v.6 Attributes

In the XML context, the word *attribute*, like some other words, has a specific technical sense. It is used to describe information that is in some sense descriptive of a specific element occurrence but not regarded as part of its content. For example, you might wish to add a *status* attribute to occurrences of some elements in a document to indicate their degree of reliability, or to add an *identifier* attribute so that you could refer to particular element occurrences from elsewhere within a document. Attributes are useful in precisely such circumstances.

[19] See ?.

Although different elements may have attributes with the same name (for example, in the TEI scheme, every element is defined as having an attribute named *n*), they are always regarded as different, and may have different values assigned to them. If an element has been defined as having attributes, the attribute values are supplied in the document instance as *attribute-value pairs* inside the start-tag for the element occurrence. An end-tag cannot contain an attribute-value specification, since it would be redundant.

The order in which attribute-value pairs are supplied inside a tag has no significance; they must, however, be separated by at least one whitespace (blank, newline, or tab) character. The value part must always be given inside matching quotation marks, either single or double[20].

For example:

```
<poem  xml:id="Poem1"  status="draft"> ... </poem>
```

Here attribute values are being specified for two attributes previously declared for the `<poem>` element: *xml:id* and *status*. For the instance of a `<poem>` in this example, represented here by ellipsis points, the *xml:id* attribute has the value P1 and the *status* attribute has the value draft. An XML processor can use the values of the attributes in any way it chooses; for example, a `<poem>` in which the *status* attribute has the value draft might be formatted differently from one in which the same attribute has the value revised; another processor might use the same attribute to determine whether or not poem elements are to be processed at all. The *xml:id* attribute is a slightly special case in that, by convention, it is always used to supply a unique value to identify a particular element occurrence, which may be used for cross-reference purposes, as discussed further below (*v.6.2 Identifiers and Indicators*).

## v.6.1 Declaring Attributes

Attributes are declared in a schema in the same way as elements. As well as specifying an attribute's name and the element to which it is to be attached, it is possible to specify (within limits) what kind of value is acceptable for an attribute.

In the compact syntax of RELAX NG, an attribute is defined by means of an attribute pattern, like the following:

```
att.status = attribute status {"draft" | "revised" | "published"}
```

This defines a new pattern, called `att.status`, whose value is an attribute pattern defining an attribute named *status*. Attribute names are subject to the same restrictions as other names in XML; they need not be unique across the whole schema, however, but only within the list of attributes for a given element.

A pattern defining the possible values for this attribute is given within the curly braces, in just the same way as a content model is given for an element pattern. In this case, the attribute's value must be one of the strings presented explicitly above.

The attribute pattern definition must be included or referenced within the definition for every element to which the attribute is attached. We therefore modify the definition for the `poem_p` pattern given above as follows:

```
poem_ p = element poem { att.status?, heading_ p?, stanza_ p+ }
```

In RELAX NG, an element pattern simply includes any attribute patterns applicable to it along with its other constituents, as shown above. Attribute patterns can also be grouped and alternated in the same way as element patterns, though this particular feature is not widely used in the TEI scheme, since it is not available to the same extent in all schema languages. Because a question mark follows the reference to the `att.status` pattern in our example, a document in which the *status* attribute is not specified will still be valid; without this occurrence indicator the *status* attribute would be required.

[20] In the unlikely event that both kinds of quotation marks are needed within the quoted string, either or both can also be presented in escaped form, using the predefined character entities `'` or `"`

Instead of supplying a list of explicit values, an attribute pattern can specify that the attribute must have a value of a particular type, for example a text string, a numeric value, a normalized date, etc. This is accomplished by supplying a pattern that refers to a *datatype*. In the example above, because a list of acceptable values is predefined, a parser can check that no `<poem>` is defined for which the *status* attribute does not have one of draft, revised, or published as its value. By contrast, with a definition such as

```
  att.status =
attribute status {text}
```

a parser would accept almost any unbroken string of characters (`status="awful"`, `status="awe-ful"`, or `status="12345678"`) as valid for this attribute. Sometimes, of course, the set of possible values cannot be predefined. Where it can, as in this case, it is generally better to do so.

Schema languages vary widely in the extent to which they support validation of attribute values. Some languages predefine a small set of possibilities. Others allow the schema designer to use values from a predefined 'library' of possible datatypes, or to add their own definitions, possibly of great complexity. A 'datatype' might be something fairly general (any positive integer), something very specific or idiosyncratic (any four-character string ending with "T"), or somewhere between the two. In the RELAX NG schemas used by the TEI, general patterns have been defined for about half a dozen datatypes (using the W3C Schema Datatype Library, and discussed further in *1.4.2. Datatype Specifications*). In addition to the two possibilities already mentioned—plain text or an explicit list of possible strings—other datatypes likely to be encountered include the following:

**boolean** values must be either true or false

**numeric** values must represent a numeric quantity of some kind

**date** values must represent a possible date and time in some calendar

Two further datatypes of particular usefulness in managing XML documents are commonly known as `ID`—for identifier—and `URI`—for Universal Resource Indicator, or pointer for short. These are discussed in the next section.

## v.6.2 Identifiers and Indicators

It is often necessary to refer to an occurrence of one textual element from within another, an obvious example being phrases such as 'see note 6' or 'as discussed in chapter 5'. When a text is being produced the actual numbers associated with the notes or chapters may not be certain. If we are using descriptive markup, such things as page or chapter numbers, being entirely matters of presentation, will not in any case be present in the marked-up text: they will be assigned by whatever processor is operating on the text (and may indeed differ in different applications). XML therefore predefines an attribute that may be used to provide any element occurrence with a special identifier, a kind of label, which may be used to refer to it from anywhere else: since it is defined in the XML namespace, the name of this attribute is *xml:id* and it is used throughout the TEI schema. Because it is intended to act as an identifier, its values must be unique within a given document. The cross-reference itself will be supplied by an element bearing an attribute of a specific kind, which must also be declared in the schema.

Suppose, for example, we wish to include a reference within the notes on one poem that refers to another poem. We will first need to provide some way of attaching a label to each poem: this is easily done using the *xml:id* attribute. Note that not every poem need carry an *xml:id* attribute and the parser may safely ignore the lack of one in those that do not. Only poems to which we intend to refer need use this attribute; for each such poem we should now include in its start-tag some unique identifier, for example:

```
<poem  xml:id="Rose"> ... </poem>
<poem  xml:id="P40"> ... </poem>
<poem> ... </poem>
```

Next we need to define a new element for the cross-reference itself. This will not have any content—it is only a pointer—but it has an attribute, the value of which will be the identifier of the element pointed at. This is achieved by the following definition:

```
  poemRef_ p = element poemRef {attribute target {anyURI}, empty}
```

The <poemRef> element has no content, but a single attribute called *target*. The value of this attribute must be a pointer or web reference of type anyURI;[21] furthermore, because there is no indication of optionality on the attribute pattern, it must be supplied on each occurrence—a <poemRef> with no referent is an impossibility.

With these declarations in force, we can now encode a reference to the poem whose *xml:id* attribute specifies that its identifier is Rose as follows:

```
Blake's poem on the sick rose
<poemRef  target="#Rose"/> ...
```

A processor may take any number of actions when it encounters a link encoded in this way: a formatter might construct an exact page and line reference for the location of the poem in the current document and insert it, or just quote the poem's title or first lines. A hypertext style processor might use this element as a signal to activate a link to the poem being referred to, for example by displaying it in a new window. Note, however, that the purpose of the XML markup is simply to indicate that a cross-reference exists: it does not necessarily determine what the processor is to do with it.

The target of a URI can be located anywhere: it may not necessarily be part of the same document, nor even located on the same computer system. Equally, it can be a resource of any kind, not necessarily an XML document or document fragment. It is thus a very convenient way of including references to non-XML data such as image files within a document. If, for example, we wished to include an illustration containing a reproduction of Blake's original in our anthology, the most appropriate method would probably be to define a new element called (for the sake of argument) <graphic> with a *target* attribute of datatype URI:

```
  graphic_ p = element graphic {att.url, empty} att.url =
attribute url {anyURI}
```

With these additions to the schema, we can now represent the location of the illustration within our text like this:

```
<poem>
  <graphic  url="http://en.wikisource.org/wiki/Image:Blake_sick_rose.jpg"/>
</poem>
```

By providing a location from which a reproduction of the required image can be downloaded, this encoding makes it possible for appropriate software able to display the image as well as record its existence.

Attributes form part of the structure of an XML document in the same way as elements, and can therefore be accessed using XPath. For example, to refer to all the poems in our anthology whose *status* attribute has the value draft, we might use an XPath such as /anthology/poem[@status='draft']. To find the headings of all such poems, we would use the XPath /anthology/poem[@status='draft']/heading.

## v.7 Other Components of an XML Document

In addition to the elements and attributes so far discussed, an XML document can contain a few other formally distinct things. An XML document may contain references to predefined strings of data that a validator must resolve before attempting to validate the document's structure; these are called *entity references*. They may be useful as a

[21] The word 'anyURI' is a predefined name, used in schema languages to mean that any *Uniform Resource Identifier* (URI) may be supplied here. The accepted syntax for URIs is an Internet Standard, defined in http://tools.ietf.org/html/rfc3986. anyURI is one of the *datatypes* defined by the W3C Schema datatype library.

means of providing 'boilerplate' text or representing character data which cannot easily be keyboarded. As noted earlier, an XML document may also contain instances of elements taken from some other *namespace*. And an XML document may also contain arbitrary signals or flags for use when the document is processed in a particular way by some class of processor (a common example in document production is the need to force a formatter to start a new page at some specific point in a document); such flags are called *processing instructions*. We discuss each of these three cases in the rest of this section.

The *XML declaration* should be supplied at the very beginning of every XML file; while it looks like a processing instruction, it is not, and may not occur anywhere other than at the very beginning of an XML file. For example:

```
<?xml version="1.0" encoding="UTF-8"?>
```

The XML declaration specifies the version number of the XML Recommendation applicable to the document it introduces (in this case, version 1.0), and optionally also the character encoding used to represent the Unicode characters within it. By default an XML document uses the character encoding UTF-8 or UTF-16; other commonly-encountered encodings include ISO 8859-1. If any character present in the document is not available in the specified character encoding, it must be represented as a character reference (*v.7.1 Character References*). The XML declaration is documentary, but should normally be supplied at the start of any XML file. If it is missing many XML-aware processors will be unable to process the associated text correctly.

## v.7.1 Character References

As mentioned above, all XML documents use the same internal character encoding. Since not all computer systems currently support this encoding directly, a special syntax is defined that can be used to represent individual characters from the Unicode character set in a portable way by providing their numeric value, in decimal or hexadecimal notation.

For example, the character *é* is represented within an XML document as the Unicode character with hexadecimal value 00E9. If such a document is being prepared on (or exported to) a system using a different character set in which this character is not available, it may instead be represented by the character reference `é` (the `x` indicating that what follows is a hexadecimal value) or `é` (its decimal equivalent). References of this type do not need to be predefined, since the underlying character encoding for XML is always the same.

To aid legibility, however, it is also possible to use a mnemonic name (such as `eacute`) for such character references, provided that each such name is mapped to the required Unicode value by means of a construct known as an *entity declaration*. A reference to a named character entity always takes the form of an ampersand, followed by the name, followed by a semicolon. For example an XML document containing the string 'T&C' might be encoded as `T&C`.

There is a small set of such character entity references that do not have to be declared because they form part of the definition of XML. These include the names used for characters such as the ampersand (`amp`) and the open angle bracket or less-than sign (`lt`), which could not easily otherwise be included in an XML document without ambiguity. Other predeclared entity names are those for quotation marks (`quot` and `apos` for double and single respectively), and for completeness the closing angle bracket or greater-than sign (`gt`).

For all other named character entities, a set of entity declarations must be provided to an XML processor before the document referring to them can be validated. The declaration itself uses a non-XML syntax inherited from SGML; for example, to define an entity named `eacute` with the replacement value é, the declaration could have any of the following forms:

```
<!ENTITY eacute "é">
```

or, using hexadecimal notation:

```
<!ENTITY eacute "&#xe9;">
```

or, using decimal notation:

```
<!ENTITY eacute "&#233;">
```

Entities of this kind are useful also for *string substitution* purposes, where the same text needs to be repeated uniformly throughout a text. For example, if a declaration such as

```
<!ENTITY TEI "Text Encoding Initiative">
```

is included with a document, then references such as `&TEI;` may be used within it, each of which will be expanded in the same way and replaced by the string 'Text Encoding Initiative' before the text is validated.

## v.7.2 Namespaces

A valid XML document necessarily specifies or implies some schema within which its constituent elements are defined and rules about their usage are given: this is a part of what it means to be 'valid'. By contrast, a well formed XML document is not required to specify a schema at all. Nonetheless, it may well be useful to indicate that the element names used in it have some defined provenance. It might also be desirable to include in a document elements that are defined (possibly differently) in different schemas. A cabinet-maker's schema might well define an element called `<table>` with very different characteristics from those of a documentalist's.

The concept of *namespace* was introduced into the XML language as a means of addressing these and related problems. If the markup of an XML document is thought of as an expression in some language, then a namespace may be thought of as analogous to the lexicon of that language. Just as a document can contain words taken from different languages, so a well-formed XML document can include elements taken from different namespaces. A namespace resembles a schema in that we may say that a given set of elements 'belongs to' a given namespace, or are 'defined by' a given schema. However, a schema is a set of element definitions, whereas a namespace is really only a property of a collection of elements: the only tangible form it takes in an XML document is its distinctive *prefix* and the identifying *name* associated with it.

Suppose for example that we wish to extend our anthology to include a complex diagram. We might start by considering whether or not to extend our simple schema to include XML markup for such features as arcs, polygons, and other graphical elements. XML can be used to represent any kind of structure, not simply text, and there are clear advantages to having our text and our diagrams all expressed in the same way.

Fortunately we do not need to invent a schema for the representation of graphical components such as diagrams; it already exists in the shape of the Scalable Vector Graphics (SVG) language defined by the W3C.[22] SVG is a widely used and rich XML vocabulary for representing all kinds of two-dimensional graphics; it is also well supported by existing software. Using an SVG-aware drawing package, we can easily draw our diagram and save it in XML format for inclusion within our anthology. When we do so, we need to indicate that this part of the document contains elements taken from the SVG namespace, if only to ensure that processing software does not confuse our `<line>` element with the SVG `<line>`, which means something quite different.

An XML document need not specify any namespace: it is then said to use the 'null' namespace. Alternatively, the root element of a document may supply a default namespace, understood to apply to all elements which have no namespace prefix. This is the function of the *xmlns* attribute which provides a unique name for the default namespace, in the form of a URI:

[22] The W3C Recommendation is defined at `https://www.w3.org/Graphics/SVG/`.

```
<anthology>
<!-- anthology markup elements here -->
</anthology>
```

In exactly the same way, on the root element for each part of our document which uses the SVG language, we might introduce the SVG namespace name:

```
<anthology>

<!-- anthology markup elements here -->
  <svg xmlns="http://www.w3.org/2000/svg">
<!-- SVG markup elements here -->
  </svg>
<!-- more anthology markup elements here -->
</anthology>
```

Although a namespace name usually uses the URI (Uniform Resource Identifier) syntax, it is not treated as an online address and an XML processor regards it just as a string, providing a longer name for the namespace.

The *xmlns* attribute can also be used to associate a short prefix name with the namespace it defines. This is very useful if we want to mingle elements from different namespaces within the same document, since the prefix can be attached to any element, overriding the implicit namespace for itself (but not its children):

```
<anthology
   xmlns:svg="http://www.w3.org/2000/svg">

<!-- anthology markup elements here -->
  <svg:svg>
<!-- SVG markup elements here -->
  </svg:svg>
<!-- more anthology markup elements here -->
</anthology>
```

There is no limit on the number of namespaces that a document can use. Provided that each is uniquely identified, an XML processor can identify those that are relevant, and validate them appropriately. To extend our example further, we might decide to add a linguistic analysis to each of the poems, using a set of elements such as `<aux>`, `<adj>`, etc., derived from some pre-existing XML vocabulary for linguistic analysis.

```
<anthology
   xmlns:gram="http://www.gram.org"
   xmlns:svg="http://www.w3.org/2000/svg">

<!-- anthology markup elements here -->
  <svg:svg>
<!-- SVG markup elements here -->
  </svg:svg>
  <line>
    <gram:itj>O</gram:itj>
    <gram:nom>Rose</gram:nom>
    <gram:pron>thou</gram:pron>
    <gram:aux>art</gram:aux>
    <gram:adj>sick</gram:adj>
  </line>
</anthology>
```

#### Marked Sections

We mentioned above that the syntax of XML requires the encoder to take special action if characters with a syntactic meaning in XML (such as the left angle bracket or ampersand) are to be used in a document to stand for themselves, rather than to signal the start of a tag or an entity reference respectively. The predefined entities `&`, `<`, and `>` provide one method of dealing with this problem, if the number of occurrences of such things is small. Other methods may be considered when the number is large, as in an XML document like the present Guidelines, which contains hundreds of examples of XML markup. One is to label the XML examples as belonging to a different namespace from that of the document itself, which is the approach taken in the present Guidelines. Another and simpler approach is provided by one of the features inherited by XML from its parent SGML: the 'marked section'.

A marked section is a block of text within an XML document introduced by the characters `<![CDATA[` and terminated by the characters `]]>`. Between these rather strange brackets, markup recognition is turned off, and any tags or entity references encountered are therefore treated as if they were plain text. For example, when we come to write the users' manual for our anthology, we may find ourselves often producing text like the following:

```
  Here is an example of the use of the <gi>line</gi> element:
<![CDATA[<line>[...]</line>]]>
```

### v.7.3 Processing Instructions

Although one of the aims of using XML is to remove any information specific to the processing of a document from the document itself, it is occasionally very convenient to be able to include such information—if only so that it can be clearly distinguished from the structure of the document. As suggested above, one common example is the need, when processing an XML document for printed output, to include a suggestion that the formatting processor might use to determine where to begin a new page of output. It is generally best to leave pagination of the output to the formatting engine alone, but there will always be occasions when it may be necessary to override these. An XML processing instruction inserted into the document is one very simple and effective way of doing this without interfering with other aspects of the markup.

Here is an example XML processing instruction:

```
  <?tex \newpage ?>
```

It begins with `<?` and ends with `?>`. In between are two strings separated by a space: the first is the name of some processor (`tex` in the above example) and the second is some data intended for the use of that processor (in this case, the instruction to start a new page). The only constraint placed by XML on the strings is that the first one must be a valid XML name other than `XML`, `Xml`, or `xml`, etc. The second string can be any arbitrary sequence of characters, not including the closing character-sequence `?>`.

## v.8 Putting It All Together

In this chapter we have discussed most of the components of an XML document and its associated schema. We have described informally how an XML document is represented, and also introduced one way of representing the rules a RELAX NG validator might use to validate it. In a working system, the following issues will also need to be addressed:

- how does a processor determine the schema (or schemas) that should be used to validate a given XML document instance?
- if a document contains entity references that must be processed before the document can be validated, where are those entities defined?
- an XML document instance may be stored in a number of different operating system files; how should they be assembled together?
- how does a processor determine which stylesheets it should use when processing an XML document, or how to interpret any processing instructions it contains?

- how does a processor enforce more exact validation than simple datatypes permit (for example of element content)?

Different schema languages and different XML processing systems take very different positions on all of these topics, since none of them is explicitly addressed in the XML specification itself. Consequently, the best answer is likely to be specific to a particular software environment and schema language. Since this chapter is concerned with XML considered independently of its processing environment, we only address them in summary detail here.

## v.8.1 Associating Entity Definitions with a Document Instance

In *v.7.1 Character References* we introduced the syntax used for the definition of named character entities such as `eacute`, which XML inherited from SGML. Different schema languages vary in the ways they make a collection of such definitions available to an XML processor, but fortunately there is one method that all current schema languages support.

As well as, and following, the XML declaration (*v.7.3 Processing Instructions*), an XML document instance may be prefixed with a special `DOCTYPE` statement. This declarative statement has been inherited by XML from SGML; in its full form it provides a large number of facilities, but we are here concerned only with the small subset of those facilities recognized by all schema languages.

Here is an example DOCTYPE statement which we might consider prefixing to the final version of our anthology:

```
  <!DOCTYPE anthology [
<!ENTITY mdash "&#2014;">
<!ENTITY legalese "This document is available under a Creative Commons
Share and Enjoy Licence">
]>
```

Any XML processor encountering this statement will use it to add the two named entities it defines to those already predefined for XML. Before the document instance itself is validated, any references to these entities will be expanded to the character string given. Thus, wherever in the document instance the string `&legalese;` appears, it will be replaced by the formulation above. This makes life a little easier for those keyboarding our anthology.[23] The word `anthology` following the string DOCTYPE in this example is, of course, the name of the root element of the document to which this declaration is prefixed; however, only an XML DTD processor will take note of this fact.

## v.8.2 Associating a Document Instance with Its Schema

In the past, different schema languages adopted entirely different attitudes to this question, leading to a variety of different methods of associating schemas with document instances. However, a W3C Working Group Note, *Associating Schemas with XML documents*, (`https://www.w3.org/TR/xml-model/`) now provides a standardized method of doing this through the use of a processing instruction:

```
  <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_
all.rng"?>
```

The *href pseudo-attribute* points to the location of the schema. This is the only mandatory pseudo-attribute, but others can be added to give more information about the schema:

```
  <?xml-model href="burgess.rng"
title="Anthony Burgess Project Schema"
schematypens="http://relaxng.org/ns/structure/1.0"
```

[23] And, indeed, for those responsible for deciding the licensing conditions if they change their minds later.

```
type="application/xml"
?>
```

See the XML Model WG Note for more information on the pseudo-attributes available and how to use them.

A document instance may be valid according to many different schemas, each appropriate to a different processing task. All of these may be expressed in the same way:

```
  <?xml-model href="tei_ tite.xsd" type="application/xml" ?>
<?xml-model href="checkNames.sch" type="application/xml" schematypens="http://purl.oclc.org/dsdl/
?>
```

This example includes a standard schema in XML Schema format, along with a Schematron schema which might be used for checking the format and linking of names.

Any modern XML processing software tool will provide convenient methods of validating documents which are appropriate to the particular schema language chosen. In the interests of maximizing portability of document instances, they should contain as little processing-specific information as possible.

### v.8.3 Assembling Multiple Resources into a Single Document

As we have already indicated, a single XML document may be made up of several different operating system files that need to be pulled together by a processor before the whole document can be validated. The XML DTD language defines a special kind of entity (a *system entity*) that can be used to embed references to whole files into a document for this purpose, in much the same way as the character or string entities discussed in *v.7.1 Character References*. Neither RELAX NG nor W3C Schema directly supports this mechanism, however, and we do not discuss it further here.

An alternative way of achieving the same effect is to use a special kind of pointer element to refer to the resources that need to be assembled, in exactly the same way as we proposed for the illustration in our anthology. The W3C Recommendation *XML Inclusions* (XInclude)[24] defines a generic mechanism for this purpose, which is supported by an increasing number of XML processors.

### v.8.4 Stylesheet Association and Processing

As mentioned above, the processing of an XML document will usually involve the use of one or more stylesheets, often but not exclusively to provide specific details of how the document should be displayed or rendered. In general, there is no reason to associate a document instance with any specific stylesheet and the schema languages we have discussed so far do not therefore make any special provision for such association. The association is made when the stylesheet processor is invoked, and is thus entirely application-specific.

However, since one very common application for XML documents is to serve them as browsable documents over the Web, the W3C has defined a procedure and a syntax for associating a document instance with its stylesheet (see `https://www.w3.org/TR/xml-stylesheet/`). This Recommendation allows a document to supply a link to a default stylesheet and also to categorize the stylesheet according to its *MIME type*, for example to indicate whether the stylesheet is written in CSS or XSLT, using a specialized form of processing instruction.

Assuming therefore that we have made a CSS-conformant stylesheet for our anthology and stored it in a file called `anthology.css` which is available from the same location as the anthology itself, we could make it available over the Web simply by adding a processing instruction like the following to the anthology:

```
  <?xml-stylesheet href="anthology.css" type="text/css"?>
```

Multiple stylesheets can be defined for the same document, and options are available to specify how a web browser should select amongst them. For example, if the document also contained a directive:

[24] `https://www.w3.org/TR/xinclude/`.

```
<?xml-stylesheet href="anthology_ m.css" type="text/css" media="mobile"?>
```

a different stylesheet called `anthology_m.css` could be used when rendering the document on a handheld device such as a mobile phone.

Most modern web browsers support CSS (although the extent of their implementation varies), and some of them support XSLT.

### Content Validation

As we noted above, most schema languages provide some degree of datatype validation for attribute values (*v.6.1 Declaring Attributes*). They vary greatly in the validation facilities they offer for the content of elements, other than the syntactic constraints already discussed. Thus, while we may very easily check that our `<stanza>` elements contain only `<line>` elements, we cannot easily check that `<line>` elements contain between five and 500 correctly-spelled English words, should we wish to constrain our poetry in such a way. Also, because attributes and elements are treated differently, it is difficult or impossible to express co-occurrence constraints: for example, if the *status* of a poem is draft we might wish to permit elements such as `<editorialQuery>` within its content, but not otherwise.

The XML DTD language offers very little beyond syntactic checking of element content. By contrast, a major impetus behind the design and development of the W3C schema language was the addition of a much more general and powerful constraint language to the existing structural constraints of XML DTDs. In RELAX NG the opposite approach was taken, in that all datatype validation, whether of attributes or element content, is regarded as external to the schema language. For attributes, as we have seen, RELAX NG makes use of the W3C Schema Datatype Library (but permits use of others). Because RELAX NG treats both elements and attributes as special cases of patterns, the same datatype validation facilities are available for element content as for attribute values; it is unlike other schema languages in this respect. In addition, for content validation, a different component of DSDL known as Schematron can be used. Schematron is a pattern matching (rather than a grammar-based) language, which allows us to test the components of a document against templates that express constraints such as those mentioned above.

Like other XML processors, Schematron uses XPath to identify parts of an XML document; in addition, it provides elements that describe assertions to be tested and conditions which must be validated, as well as elements to report the results of the test.

# vi Languages and Character Sets

The documents which users of these Guidelines may wish to encode encompass all kinds of material, potentially expressed in the full range of written and spoken human languages, including the extinct, the non-existent, and the conjectural. Because of this wide scope, special attention has been paid to two particular aspects of the representation of linguistic information often taken for granted: language identification and character encoding.

Even within a single document, material in many different languages may be encountered. Human culture, and the texts which embody it, is intrinsically multilingual, and shows no sign of ceasing to be so. Traditional philologists and modern computational linguists alike work in a polyglot world, in which code-switching (in the linguistic sense) and accurate representation of differing language systems constitute the norm, not the exception. The current increased interest in studies of linguistic diversity, most notably in the recording and documentation of endangered languages, is one aspect of this long standing tradition. Because of their historical importance, the needs of endangered and even extinct languages must be taken into account when formulating Guidelines and recommendations such as these.

Beyond the sheer number and diversity of human languages, it should be remembered that in their written forms they may deploy a huge variety of scripts or writing systems. These scripts are in turn composed of smaller units, which for simplicity we term here characters. A primary goal when encoding a text should be to capture enough information for subsequent users to correctly identify not only the constituent characters, but also the language and script. In this chapter we address this requirement, and propose recommended mechanisms to indicate the languages, scripts and characters used in a document or a part thereof.

Identification of language is dealt with in *vi.1 Language Identification*. In summary, it recommends the use of pre-defined identifiers for a language where these are available, as they increasingly are, in part as a result of the twin pressures of an increasing demand for language-specific software and an increased interest in language documentation. Where such identifiers are not available or not standardized, these Guidelines recommend a method for documenting language identifiers and their significance, in the same way as other metadata is documented in the TEI header.

Standardization of the means available to represent characters and scripts has moved on considerably since the publication of the first version of these Guidelines. At that time, it was essential to explicitly document the characters and encoded character sets used by almost any digital resource if it was to have any chance of being usable across different computer platforms or environments, but this is no longer the case. With the availability of the Unicode standard, more than 128,000 different characters representing almost all of the world's current writing systems are available and usable in any XML processing environment without formality. Nevertheless, however large the number of standardized characters, there will always be a need to encode documents which use non-standard characters and glyphs, particularly but not exclusively in historical material. The second part of this chapter discusses in some detail the concepts and practice underlying this standard, and also introduces the methods available for extending beyond it, which are more fully discussed in *5. Characters, Glyphs, and Writing Modes*.

## vi.1 Language Identification

Identification of the language a document or part thereof is written in is a crucial requirement for many envisioned usages of an electronic document. The TEI therefore accommodates this need in the following way:

- A global attribute *xml:lang* is defined for all TEI elements. Its value identifies the language and writing system used.
- The TEI header has a section set aside for the information about the languages used in a document: see further *2.4.2. Language Usage*.

The value of the attribute *xml:lang* identifies the language (and, optionally, script) using a coded value. For maximal compatibility with existing processes, the identifier for the language must be constructed as in *Best Current Practice 47*[25]. This *same* identifier has to be used to identify the corresponding <language> element in the TEI header, if one is present.

The first part of BCP 47 is called *Tags for Identifying Languages*, and proposes the following mechanism for constructing an identifier (tag) for languages as administered by the Internet Assigned Numbers Authority (IANA). The tag is assembled from a sequence of subtags separated by the hyphen (-, U+002D) character. It gives the language

[25] Currently BCP 47 comprises two Internet Engineering Task Force documents, referred to separately as RFC 5646 and RFC 4647; over time, other IETF documents may succeed these as the best current practice.

(possibly further identified with a sublanguage), a script, and a region for the language, each possibly followed by a variant subtag.

The authoritative list of registered subtags is maintained by IANA and is available at `https://www.iana.org/assignments/language-subtag-registry`. For a good general overview of the construction of language tags, see `https://www.w3.org/International/articles/language-tags/`, and for a practical step-by-step guide, see `https://www.w3.org/International/questions/qa-choosing-language-tags.en.php`.

In addition to the list of registered subtags, BCP 47 provides extensions that can be employed by private convention. The constructs provided can thus be used to generate identifiers for any language, past and present, in any usage in any area of the world. If such private extensions are used within the context of the TEI, they should be documented within the <language> element of the TEI header, which might also provide a prose description of the language described by the language tag.

While language, region, and script can be adequately identified using this mechanism, there is only very rough provision to express a dimension of time for the language of a document; those codes provided (e.g. `grc` for 'Greek, Ancient (to 1453)') might not reflect the segments appropriate for a text at hand. Text encoders might express the time window of the language used in the document by means of the extension mechanism defined in BCP 47 and relate that to a <date> element in the corresponding <language> section of the TEI header.

Equivalences to language identifiers by other authorities can be given in the <language> section as well, but no formal mechanism for doing so has been defined.

The scope of the language identification extends to the whole subtree of the document anchored at the element that carries the *xml:lang* attribute, including all elements and those attributes, if any, where a language might apply.[26]

## vi.2 Characters and Character Sets

All document encoding has to do with representing one thing by another in an agreed and systematic way. Applied to the smallest distinctive units in any given writing system, which for the moment we may loosely call 'characters', such representation raises surprisingly complex and troublesome issues. The reasons are partly historical and partly to do with conceptual unclarities about what is involved in identifying, encoding, processing and rendering the characters of a natural language.

### vi.2.1 Historical Considerations

When the first methods of representing text for storage or transmission by machines were devised, long before the development of computers, the overriding aim was to identify the smallest set of symbols needed to convey the essential semantic content, and to encode that symbol set in the most economical way that the storage or transmission media allowed. The initial outcome were systems that encoded only such content as could be expressed in uppercase letters in the Latin script, plus a few punctuation marks and some 'control characters' needed to regulate the storage and transmission devices. Such encodings, originally developed for telegraphy, strongly influenced the way the pioneers of computing conceived of and implemented the handling of text, with consequences that are with us still.

For many years after the invention of computers, the way they represented text continued to be constrained by the imperative to use expensive resources with maximal efficiency. Even when storage and processing costs began their dramatic fall, the Anglo-centric outlook of most hardware designers and software engineers hampered initiatives to devise a more generous and flexible model for text representation. The wish to retain compatibility with 'legacy' data was an additional disincentive. Eventually, tension in East Asia between commitment to technological progress and the inability of existing computers to cope with local writing systems led to decisive developments. Japanese, Korean, and Chinese standards bodies, who long before the advent of computers had been engaged in the specification of character sets, joined with computer manufacturers and software houses to devise ways of mapping those character sets to numeric encodings and processing the resulting text data.

Unfortunately, in the early years there was little or no co-ordination among either the national standards bodies or the manufacturers concerned, so that although commercial necessity dictated that these various local standards were all compatible with the representation of US-American English, they were not straightforwardly compatible with one another. Even within Japan itself there emerged a number of mutually incompatible systems, thanks to a mixture of commercial rivalry, disagreements about how best to manage certain intractable problems, and the fact

[26] This excludes all attributes where a non-textual datatype has been specified, for example tokens, boolean values, dates, and predefined value lists.

that such pioneering work inevitably involved some false starts, leading to incompatibilities even between successive products of the same bodies. Roughly at the same time, and for similar reasons, multiple and incompatible ways of representing languages that use Cyrillic scripts were devised, along with methods of encoding ancient writing systems which inevitably could not aim for compatibility with other writing systems apart from basic Latin script. Many of the earliest projects that fed into the TEI were shaped in this developmental phase of the computerized representation of texts, and it was also the context in which SGML was devised and finalized.

SGML had of necessity to offer ways of coping with multiple writing systems in multiple representations; or rather, it provided a framework within which SGML-compliant applications capable of handling such multiple representations might be developed by those with sufficient financial and personnel resources (such as are seldom found in academia). Earlier editions of these Guidelines offered advice on character set and writing system issues addressed to the condition of those for whom SGML was the only feasible option. That advice is here substantially altered because of two closely-related developments: the availability of the ISO/Unicode character set as an international standard, and the emergence of XML and related technologies which are committed to the theory and practice of character representation which Unicode embodies.

### vi.2.2 Terminology and Key Concepts

Before the significance of Unicode and the implications of the association between XML and Unicode can be adequately explained, it is necessary to clarify some key concepts and attempt to establish an adequately precise terminology for them.

Figure 1: Examples of the latin *a*, in both lower and upper case, rendered with different fonts.

The word 'character' will not of itself take us very far towards greater terminological precision. It tends to be used to refer indiscriminately both to the visible symbol on a page and to the letter or ideograph which that symbol represents, two things that it is essential to keep conceptually distinct. The visible symbol obviously has some aspects by which we interpret it as representing one character rather than another; but its appearance may also be significantly determined by features that have no effect on our notion of which character in a writing system it represents. A familiar instance is the lowercase *a*, which in printed texts may be represented either by a 'single storey' symbol (cf. figure 1 in the examples from URW Gothic L on the bottom row) or by a 'two storey' version (as in figure 1 in the examples from Umpush, or URW Bookman L Demi Bold). We say that the single and double-storey symbols both represent one and the same the same *abstract character a* using two different *glyphs*. Similarly, an uppercase *A* in a serif typeface has additional strokes that are absent from the same letter when printed using a sans-serif typeface, so that once again we have differing glyphs standing for the same abstract character. The distinction between abstract characters and glyphs is fundamental to all machine processing of documents.

In most scholarly encoding projects, the accurate recording of the abstract characters which make up the text is of prime importance, because it is the essential prerequisite of digitizing and processing the document without semantic loss. In many cases (though there are important exceptions, to be touched on shortly) it may not be necessary to

encode the specific glyphs used to render those abstract characters in the original document. An encoding that faithfully registers the abstract characters of a document allows us to search and analyse our document's content, language, and structure, and to access its full semantics. That same encoding, however, may not contain sufficient information to allow an exact visual representation of the glyphs in the source text or manuscript to be recreated.

The importance of this distinction between information content and its visual representation is not always immediately apparent to people unused to the specific complexities of text handling by machine. Such users tend to ask first what (in order of conceptual priority) should actually be their very last question: how do I get a physical image that looks like character x in my source document to appear on to the screen or the output page? Their first question should in fact be: how can I get an abstract representation of character x into my encoded document in a way that will be universally and unambiguously identifiable, no matter what it happens to look like in printout or on any particular display? And occasionally the response they receive as a result of their misguided initial question is a custom 'solution' that satisfies their immediate rendering wishes at the price of making their underlying document unintelligible to other users (or even to the original user in other times and places) because it encodes the abstract character in an idiosyncratic way.

That said, there will certainly be documents or projects where it is a matter of scholarly significance that the compositor or scribe chose to represent a given abstract character using one particular glyph or set of strokes rather than a semantically-equivalent but visually distinct alternative, and in that case the specific appearance of the form will have to be encoded in one way or another. But that encoding need not (and in most cases will not) involve a notation that visually resembles the original, any more than italicized text in an original document will be represented by the use of italic characters in the encoded version.

A collection of the abstract characters needed to represent documents in a given writing system is known as a *character set*, and the character set or *character repertoire* of a processing or rendering device is the set of abstract characters that it is equipped to recognize and handle appropriately. There is, however, a subtle distinction between these two parallel uses of the same term, involving one more key concept which it is essential to grasp. The character set of a document (or the writing system in which it is recorded) is purely a collection of abstract characters. But the character set of a computing device is a set of abstract characters which have been mapped in a well-defined way to a set of numbers or *code points* by which the device represents those abstract characters internally. It can therefore be referred to as a *coded character set*, meaning a set of abstract characters each of which has been assigned a numerical code point (or in some instances a sequence of code points) which unambiguously identifies the character concerned.

It is now possible to use this terminology to say what Unicode is: it is a coded character set, devised and actively maintained by an international public body, where each abstract character is identified by a unique name and assigned a distinctive code point.[27] Unicode is distinguished from other coded character sets by its (current and potential) size and scope; its built-in provision for (in practical terms) limitless expansion; the range and quality of linguistic and computational expertise on which it draws; the stability, authority, and accessibility it derives from its status as an international public standard; and, perhaps most importantly, the fact that today it is implemented by almost every provider of hardware and software platforms worldwide.

### vi.2.3 Abstract Characters, Glyphs, and Encoding Scheme Design

The distinction between abstract characters and glyphs can be crucial when devising an encoding scheme. When performing searches, text retrieval, or creating concordances, users of electronic text will expect the system to recognize and treat different glyphs as instances of the same character; but when perusing the text itself they may well expect to see glyph variants preserved and rendered. When encoding a pre-existing text, the encoder should determine whether a particular letter or symbol is a character or a glyphic variant. The Unicode Consortium and an ISO work group (ISO/IEC JTC1 SC2/WG2) have developed a detailed model of the relationship between characters and glyphs. This model, presented in Unicode Technical Report 17: Character Encoding Model, is the underpinning of much standards work since, including the current chapter.

The model makes explicit the distinction between two different properties of the components of written language:

- their content, i.e. its meaning and phonetic value (represented by characters)
- their graphical appearance (represented by glyphs).

[27] Although only Unicode is mentioned here explicitly, it should be noted that the character repertoire and assigned code points of Unicode and the ISO standard 10646 are identical and maintained in a way that ensures this continues to be the case.

When searching for information, a system generally operates on the content aspects of characters, with little or no attention to their appearance. A layout or formatting process, on the other hand, must of necessity be concerned with the exact appearance of characters. Of course, some operations (hyphenation for example) require attention to both kinds of feature, but in general the kind of text encoding described in these Guidelines tends to focus on content rather than appearance (see further *3.3. Highlighting and Quotation*).

An encoder wishing to record information about which glyphs are present in a given document may do so at either or both of two levels:

- the level of character encoding, using an appropriate Unicode code point to represent the glyph concerned
- the markup level, with the glyph indicated via appropriate elements or attributes

The encoding practice adopted may be guided by, among other things, an assessment of the most frequent uses to which the encoded text will be put. For example, if recognition of identical characters represented by a variety of glyphs is the main priority, it may be advisable to represent the glyph variations at markup level, so that the character value can be immediately exposed to the indexing and retrieval software. Plainly, an encoding project will need to consider such issues carefully and document the outcome of their deliberations in their TEI customization file (or other local encoding documentation) to ensure encoding consistency. Using Unicode code points to represent glyph information requires that such choices be documented in the TEI header. Such documentation cannot of itself guarantee proper display of the desired glyph but at least makes the intention of the encoder discoverable.

At present the Unicode Standard does not offer detailed specifications for the encoding of glyph variations. These Guidelines do give some recommendations; some discussion of related matters is given in *12. Representation of Primary Sources*, and *5. Characters, Glyphs, and Writing Modes* offers some features for the definition of variant glyphs.

## vi.2.4 Entry of Characters

The entry of characters was much more complicated before the near-universal adoption of Unicode, for which there are *Input Method Editors* (IMEs) available in most languages and fonts that provide glyphs for the full range of the Unicode specification. In those rare situations where there is difficulty entering the specific character you want, or some problem representing it on the system you are working in, *Numeric Character References* (NCRs) should be used. These take the general form `&#D;` where `D` is an integer representing the code point of the character in base 10, or `&#xH;`, where `H` is the code point in hexadecimal notation. Every XML processor is capable of recognising NCRs and replacing them with the required code point value without needing access to any additional data. The disadvantage of NCRs as a means of entering, representing and proofing character data is that most human beings find them anything but 'readable' and it is all too easy for the wrong character to be entered in error and retained undetected. Where characters are not defined in Unicode, these Guidelines provide advice on the strategies available for handling their representation in Chapter 25 Representation of non-standard Characters and Glyphs.

## vi.2.5 Output of Characters

The rendering of the encoded text is a complicated process that depends largely on the purpose, external requirements, local equipment and so forth, it is thus outside the scope of coverage for these Guidelines.

It might nevertheless be helpful to put some of the terminology used for the rendering process in the context of the discussion of this chapter. As was mentioned above, Unicode encodes abstract characters, not specific glyphs. For any process that makes characters visible, however, concrete, specifically designed glyph shapes have to be used. For a printing process, for example, these shapes describe exactly at which point ink has to be put on the paper and which areas have to be left blank. If we want to print a character from the Latin script, besides the selection of the overall glyph shape, this process also requires that a specific weight and size of the font has been selected, and to what degree the shape should be slanted. Beyond individual characters, the overall typesetting process also follows specific rules for calculating the distance between characters, for determining how much whitespace occurs between any two words, and how long each line should be (and thus at which points a new line begins), and so forth.

If we concern ourselves only with the rendering process of the characters themselves, leaving out all these other parameters, we will realize that of all the information required for this process, only a small amount will be drawn from the encoded text itself. This information is the code point used to encode the character in the document. With this information, the font selected for printing will be queried to provide a glyph shape for this character. Some modern font formats (e.g. OpenType) implement a sophisticated mapping from a code point to the glyph selected, which might take into account surrounding characters (to create ligatures where necessary) and the language or

even area this character is printed for to accommodate different typesetting traditions and differences in the usage of glyphs.

A TEI document might provide some of the information that is required for this process, for example by identifying the linguistic context with the *xml:lang* attribute. The selection of fonts and sizes is usually done in a stylesheet, while the actual layout of a page is determined by the typesetting system used. Similarly, if a document is rendered for publication on the Web, information of this kind can be shipped with the document in a stylesheet.[28]

## vi.2.6 Unicode and XML

XML was designed with Unicode in mind as its means of representing abstract characters. It is possible to use other character encoding schemes, but in general they are best avoided, as you run the risk of encountering compatibility issues with different XML processors, as well as potential difficulties with rendering their output. We recommend using the *UTF-8* encoding, which for the Basic Latin range is identical to ASCII, and which uses a variable-length set of bytes to represent characters. It should be noted that it is not sufficient simply to declare in the XML Declaration that a document is in UTF-8 format. Doing so merely means that processors will treat the content therein as if it were UTF-8, and may fail to process the document if it is not. For further discussion of UTF-8, see the section below on *vi.2.8 Issues Arising from the Internal Representations of Unicode*.

## vi.2.7 Special Aspects of Unicode Character Definitions

### Compatibility Characters

The principles of Unicode are judiciously tempered with pragmatism. This means, among other things, that the actual repertoire of characters which the standard encodes, especially those parts dating from its earlier days, include a number of items which on a strict interpretation of the Unicode Consortium's theoretical approach should not have been regarded as abstract characters in their own right. Some of these characters are grouped together into a code-point regions assigned to *compatibility characters*. Ligatures are a case in point. Ligatures (e.g. the joining of adjacent lowercase letters 's' and 't' or 'f' and 'i' in Latin scripts, whether produced by a scribal practice of not lifting the pen between strokes or dictated by the aesthetics of a type design) are representational features with no added semantic value beyond that of the two letters they unite (though for historians of typography their presence and form in a given edition may be of scholarly significance). However, by the time the Unicode standard was first being debated, it had become common practice to include single glyphs representing the more common ligatures in the repertoires of some typesetting devices and high-end printers, and for the coded character sets built into those devices to use a single code point for such glyphs, even though they represent two distinct abstract characters. So as to increase the acceptance of Unicode among the makers and users of such devices, it was agreed that some such pseudo-characters should be incorporated into the standard as compatibility characters. Nevertheless, if a project requires the presence of such ligatured forms to be encoded, this should normally be done via markup, not by the use of a compatibility character. That way, the presence of the ligature can still be identified (and, if desired, rendered visually) where appropriate, but indexing and retrieval software will treat the code points in the document as a simple sequential occurrence of the two constituent characters concerned and so correctly align their semantics with non-ligatured equivalents. Such ligatures should not be confused with digraphs (usually) indicating diphthongs, as in the French word "cœur". A digraph is an atomic orthographic unit representing an abstract character in its own right, not purely an amalgamation of glyphs, and indexing and retrieval software will need to treat it as such. Where a digraph occurs in a source text, it should normally be encoded using the appropriate code point for the single abstract character which it represents.

### Precomposed and Combining Characters and Normalization

The treatment of characters with diacritical marks within Unicode shows a similar combination of rigour and pragmatism. It is obvious enough that it would be feasible to represent many characters with diacritical marks in Latin and some other scripts by a sequence of code points, where one code point designated the base character and the remainder represented one or more diacritical marks that were to be combined with the base character to produce an appropriate glyphic rendering of the abstract character concerned. From its earliest phase, the Unicode Consortium espoused this view in theory but was prepared in practice to compromise by assigning single code points to *precomposed* characters which were already commonly assigned a single distinctive code point in existing encoding schemes. This means, however, that for quite a large number of commonly-occurring abstract characters,

[28] The World Wide Web Consortium provides recommendations for two standard stylesheet languages: either CSS or XSL could be used for this purpose.

Unicode has two different, but logically and semantically equivalent encodings: a *precomposed* single code point, and a code point sequence of a base character plus one or more *combining* diacritics. Scripts more recently added to Unicode no longer exhibit this code-point duplication (in current practice no new precomposed characters are defined where the use of combining characters is possible) but this does nothing to remove the problem caused by the duplications from older character sets that have been permanently embodied in Unicode. Together with essentially analogous issues arising from the encoding of certain East Asian ideographs. This duplication gives rise to the need to practice *normalization* of Unicode documents. Normalization is the process of ensuring that a given abstract character is represented in one way only in a given Unicode document or document collection. The Unicode Consortium provides four standard normalization forms, of which the *Normalization Form C* (NFC) seems to be most appropriate for text encoding projects. The NFC, as far as possible, defines conversions for all base characters followed by one or more combining characters into the corresponding precomposed characters. The World Wide Web Consortium has produced a document entitled *Character Model for the World Wide Web 1.0*[29], which among other things discusses normalization issues and outlines some relevant principles. An authoritative reference is Unicode Standard Annex #15 *Unicode Normalization Forms*[30].

It is important that every Unicode-based project should agree on, consistently implement, and fully document a comprehensive and coherent normalization practice. As well as ensuring data integrity within a given project, a consistently implemented and properly documented normalization policy is essential for successful document interchange. While different input methods may themselves differ in what normalization form they use, any programming language that implements Unicode will provide mechanisms for converting between normalization forms, so it is easy in practice to ensure that all documents in a project are in a consistent form, even if different methods are used to enter data.

#### Character Semantics

In addition to the Universal Character Set itself, the Unicode Consortium maintains a database of additional character semantics[31]. This includes names for each character code point and normative properties for it. Character properties, as given in this database, determine the semantics and thus the intended use of a code point or character. The database also contains information that might be needed for correctly processing this character for different purposes. It is an important reference in determining which Unicode code point to use to encode a certain character.

In addition to the printed documentation and lists made available by the Unicode consortium, the information it contains may also be accessed by a number of search systems over the Web (e.g. `http://www.eki.ee/letter/`). Examples of character properties included in the database include case, numeric value, directionality, and, (where applicable) status as a 'compatibility character'[32]. Where a project undertakes local definition of characters with code points in the PUA, it is desirable that any relevant additional information about the characters concerned should be recorded in an analogous way, as further discussed under *5. Characters, Glyphs, and Writing Modes*.

### vi.2.8 Issues Arising from the Internal Representations of Unicode

In theory it should not be necessary for encoders to have any knowledge of the various ways in which Unicode code points can be represented internally within a document or in the memory of a processing system, but experience shows that problems frequently arise in this area because of mistaken practice or defective software, and in order to recognize the resulting symptoms and correct their causes an outline knowledge of certain aspects of Unicode internal representation is desirable. There are three encodings of Unicode available for use: UTF-8, which uses 1–4 bytes per character, UTF-16, which uses 2–4, and UTF-32, which uses 4 bytes per character. Current practice for documents to be transmitted via the Web recommends only UTF-8.[33]

#### Encoding Errors Related to UTF-8

The code points assigned by Unicode 3.0 and later are notionally 32-bit integers, and the most straightforward way to represent each such integer in computer storage would be to use 4 eight-bit bytes. However, many of the code

[29] Available at `https://www.w3.org/TR/charmod/`.

[30] available at `https://www.unicode.org/reports/tr15/`

[31] `https://www.unicode.org/ucd/`

[32] For further details, see *The Unicode Character Property Model* (Unicode Technical Report #23), at `https://www.unicode.org/reports/tr23/`.

[33] See the W3C Internationalization document, *Choosing & applying a character encoding* at `https://www.w3.org/International/questions/qa-choosing-encodings`

points for characters most commonly used in Latin scripts can be represented in one byte only and the vast majority of the remainder which are in common use (including those assigned from the most frequently used PUA range) can be expressed in two bytes alone. This accounts for the use of UTF-8 and UTF-16 and their special place in the XML standard. UTF-8 and UTF-16 are ways of representing 32-bit code points in an economical way.

UTF-8 is a variable length encoding: the more significant bits there are in the underlying code point (or in everyday terminology the bigger the number used to represent the character), the more bytes UTF-8 uses to encode it. What makes UTF-8 particularly attractive for representing Latin scripts, explaining its status as the default encoding in XML documents, is that all code points that can be expressed in seven or fewer bits (the 127 values in the original ASCII character set) are also encoded as the same seven or fewer bits (and therefore in a single byte) in UTF-8. That is why a document which is actually encoded in pure 7-bit ASCII can be fed to an XML processor without alteration and without its encoding being explicitly declared: the processor will regard it as being in the UTF-8 representation of Unicode and be able to handle it correctly on that basis.

However, even within the domain of Latin-based scripts, some projects have documents which use characters from 8 bit extensions to ASCII, e.g. those in the ISO-8859-n series of encodings, and the way characters which under ISO-8859-n use all eight bits are encoded in UTF-8 is significantly different, giving rise to puzzling errors. Abstract characters that have a *single* byte code point where the highest bit is set (that is, they have a decimal numeric representation between 129 and 255) are encoded in ISO-8859-n as a *single* byte with the same value as the code point. But in UTF-8 code-point values inside that range are expressed as a *two* byte sequence. That is to say, the abstract character in question is no longer represented in the file or in memory by the same number as its code-point value: it is *transformed* (hence the T in UTF) into a sequence of two different numbers. Now as a side-effect of the way such UTF-8 sequences are derived from the underlying code-point value, many of the single-byte eight-bit values employed in ISO-8859-n encodings are illegal in UTF-8.

This complicated situation has a simple consequence which can cause great bewilderment. XML processors will effortlessly handle character data in pure 7-bit ASCII without that encoding needing to be declared to the parser, and will similarly accept documents encoded in an undeclared ISO-8859-n encoding if they happen to use no characters outside the strict ASCII subset of the ISO character sets; but the parse will immediately fail if an eight-bit character from an ISO-8859-n set is encountered in the input stream, unless the document's encoding has been explicitly and correctly declared. Explicitly declaring the encoding ought to solve the problem, and if the file is correctly encoded throughout, it will do so. But projects dealing with documents of sufficient age may find that they have to deal with some files encoded in UTF-8 along with others in, say, ISO-8859-1. Such encoding differences may go unnoticed, especially if the proportion of characters where the internal encodings are distinguishable is relatively small (for example in a long English text with a smattering of French words). These types of error may or may not manifest in actual processing errors, and may only become visible as 'garbage' characters in the eventual display of documents.

In projects that routinely handle documents in non-Latin scripts, everyone is well aware of the need to ensure correct and consistent encoding, so in such places mixed encoding problems seldom arise, and when they do are readily identified and remedied. Real confusion tends to arise, however, in projects which have a low awareness of the issues because they employ predominantly unaccented Latin characters, with only thinly-distributed instances of accented letters, or other 'special characters' where the internal representation under ISO-8859-n and UTF-8 are different (such as the copyright symbol, or, a frequent troublemaker where eventual HTML output is envisaged, the 'non-breaking space'). Even, or especially, if such projects view themselves as concerned only with English documents, the close relationship between XML and Unicode means they will need to acquire an understanding of these encoding issues and develop procedures which assure consistency and integrity of encoding and its correct declaration, including the use of appropriate software for transcoding and verification.

# 1 The TEI Infrastructure

This chapter describes the infrastructure for the encoding scheme defined by these Guidelines. It introduces the conceptual framework within which the following chapters are to be understood, and the means by which that conceptual framework is implemented. It assumes some familiarity with XML and XML schemas (see chapter *v A Gentle Introduction to XML*) but is intended to be accessible to any user of these Guidelines. Other chapters supply further technical details, in particular chapter *23. Documentation Elements* which describes the XML schema used to express these Guidelines themselves, and chapter *24. Using the TEI* which combines a discussion of modification and conformance issues with a description of the intended behaviour of an ODD processor; these chapters should be read by anyone intending to implement a new TEI-based system.

The TEI encoding scheme consists of a number of *modules*, each of which declares particular XML elements and their attributes. Part of an element's declaration includes its assignment to one or more element *classes*. Another part defines its possible content and attributes with reference to these classes. This indirection gives the TEI system much of its strength and its flexibility. Elements may be combined more or less freely to form a *schema* appropriate to a particular set of requirements. It is also easy to add new elements which reference existing classes or elements to a schema, as it is to exclude some of the elements provided by any module included in a schema.

In principle, a TEI schema may be constructed using any combination of modules. However, certain TEI modules are of particular importance, and should always be included in all but exceptional circumstances: the module `tei` described in the present chapter is of this kind because it defines classes, macros, and datatypes which are used by all other modules. The `core` module, defined in chapter *3. Elements Available in All TEI Documents* contains declarations for elements and attributes which are likely to be needed in almost any kind of document, and is therefore recommended for global use. The `header` module defined in chapter *2. The TEI Header* provides declarations for the metadata elements and attributes constituting the TEI header, a component which is required for TEI conformance, while the `textstructure` module defined in chapter *4. Default Text Structure* declares basic structural elements needed for the encoding of most book-like objects. Most schemas will therefore need to include these four modules.

The specification for a TEI schema is itself a TEI document, using elements from the module described in chapter *23. Documentation Elements*: we refer to such a document informally as an *ODD* document, from the design goal originally formulated for the system: 'One Document Does it all'. Stylesheets for maintaining and processing ODD documents are maintained by the TEI, and these Guidelines are also maintained as such a document. As further discussed in *24.5. Implementation of an ODD System*, an ODD document can be processed to generate a schema expressed using any of the three schema languages currently in wide use: the XML DTD language, the ISO RELAX NG language, or the W3C Schema language, as well as to generate documentation such as the *Guidelines* and their associated web site.

The bulk of this chapter describes the TEI infrastructure module itself. Although it may be skipped at a first reading, an understanding of the topics addressed here is essential for anyone planning to take full advantage of the TEI customization techniques described in chapter *24.3. Customization*.

The chapter begins by briefly characterizing each of the modules available in the TEI scheme. Section *1.2. Defining a TEI Schema* describes in general terms the method of constructing a TEI schema in a specific schema language such as XML DTD language, RELAX NG, or W3C Schema.

The next and largest part of the chapter introduces the attribute and element classes used to define groups of elements and their characteristics (section *1.3. The TEI Class System*).

Finally, section *1.4. Macros* introduces the concept of *macros*, which are used to express some commonly used content models, and lists the *datatypes* used to constrain the range of legal values for TEI attributes (section *1.4.2. Datatype Specifications*).

## 1.1 TEI Modules

These Guidelines define several hundred elements and attributes for marking up documents of any kind. Each definition has the following components:

- a prose description
- a formal declaration, expressed using a special-purpose XML vocabulary defined by these Guidelines in combination with elements taken from the ISO schema language RELAX NG
- usage examples

Each chapter of these Guidelines presents a group of related elements, and also defines a corresponding set of declarations, which we call a *module*. All the definitions are collected together in the reference sections provided as an appendix. Formal declarations for a given chapter are collected together within the corresponding module. For convenience, each element is assigned to a single module, typically for use in some specific application area, or to support a particular kind of usage. A module is thus simply a convenient way of grouping together a number of associated element declarations. In the simple case, a TEI schema is made by combining together a small number of modules, as further described in section *1.2. Defining a TEI Schema* below.

The following table lists the modules defined by the current release of these Guidelines:

| Module name | Where defined | Formal public identifier |
|---|---|---|
| tei | *1. The TEI Infrastructure* | TEI Infrastructure |
| header | *2. The TEI Header* | Common Metadata |
| core | *3. Elements Available in All TEI Documents* | Common Core |
| textstructure | *4. Default Text Structure* | Default Text Structure |
| gaiji | *5. Characters, Glyphs, and Writing Modes* | Character and Glyph Documentation |
| verse | *6. Verse* | Verse |
| drama | *7. Performance Texts* | Performance Texts |
| spoken | *8. Transcriptions of Speech* | Transcribed Speech |
| cmc | *9. Computer-mediated Communication* | TEI-CMC |
| dictionaries | *10. Dictionaries* | Dictionaries |
| msdescription | *11. Manuscript Description* | Manuscript Description |
| transcr | *12. Representation of Primary Sources* | Transcription of Primary Sources |
| textcrit | *13. Critical Apparatus* | Text Criticism |
| namesdates | *14. Names, Dates, People, and Places* | Names, dates, persons and places |
| figures | *15. Tables, Formulæ, Graphics, and Notated Music* | Tables, Formulæ, Notated Music, Figures |
| corpus | *16. Language Corpora* | Metadata for Language Corpora |
| linking | *17. Linking, Segmentation, and Alignment* | Linking, Segmentation, and Alignment |
| analysis | *18. Simple Analytic Mechanisms* | Analysis and Interpretation |
| iso-fs | *19. Feature Structures* | Feature Structures |
| nets | *20. Graphs, Networks, and Trees* | Graphs, networks, and trees |
| certainty | *22. Certainty, Precision, and Responsibility* | Certainty and Uncertainty |
| tagdocs | *23. Documentation Elements* | Documentation Elements |

For each module listed above, the corresponding chapter gives a full description of the classes, elements, and macros which it makes available when it is included in a schema. Other chapters of these Guidelines explore other aspects of using the TEI scheme.

## 1.2 Defining a TEI Schema

To determine that an XML document is valid (as opposed to merely well-formed), its structure must be checked against a schema, as discussed in chapter *v A Gentle Introduction to XML*. For a valid TEI document, this schema must be a conformant TEI schema, as further defined in chapter *24.4. Conformance*. Local systems may allow their schema to be implicit, but for interchange purposes the schema associated with a document *must* be made explicit. The method of doing this recommended by these Guidelines is to provide explicitly or by reference a TEI schema specification against which the document may be validated.

A TEI-conformant schema is a specific combination of TEI modules, possibly also including additional declarations that modify the element and attribute declarations contained by each module, for example to suppress or rename some elements. The TEI provides an application-independent way of specifying a TEI schema by means of the <schemaSpec> element defined in chapter *23. Documentation Elements*. The same system may also be used to specify a schema which extends the TEI by adding new elements explicitly, or by reference to other XML vocabularies. In either case, the specification may be processed to generate a formal schema, expressed in a variety of specific schema languages, such as XML DTD language, RELAX NG, or W3C Schema. These output schemas can then be used by an XML processor such as a validator or editor to validate or otherwise process documents. Further information about the processing of a TEI formal specification is given in chapter *24. Using the TEI*.

### 1.2.1 A Simple Customization

The simplest customization of the TEI scheme combines just the four recommended modules mentioned above. In ODD format, this schema specification takes this form:

```
<schemaSpec ident="TEI-minimal" start="TEI">
  <moduleRef key="tei"/>
  <moduleRef key="header"/>
  <moduleRef key="core"/>
  <moduleRef key="textstructure"/>
</schemaSpec>
```

This schema specification contains references to each of four modules, identified by the *key* attribute on the <moduleRef> element. The schema specification itself is also given an identifier (`TEI-minimal`). An ODD processor will generate an appropriate schema from this set of declarations, expressed using the XML DTD language, the ISO RELAX NG language, the W3C Schema language, or in principle any other adequately powerful schema language. The resulting schema may then be associated with the document instance by one of a number of different mechanisms, as further described in chapter *v A Gentle Introduction to XML*. The start point (or root element) of document instances to be validated against the schema is specified by means of the *start* attribute. Further information about the processing of an ODD specification is given in *24.5. Implementation of an ODD System*.

### 1.2.2 A Larger Customization

These Guidelines introduce each of the modules making up the TEI scheme one by one, and therefore, for clarity of exposition, each chapter focusses on elements drawn from a single module. In reality, of course, the markup of a text will draw on elements taken from many different modules, partly because texts are heterogeneous objects, and partly because encoders have different goals. Some examples of this heterogeneity include:

- a text may be a collection of other texts of different types: for example, an anthology of prose, verse, and drama;
- a text may contain other smaller, embedded texts: for example, a poem or song included in a prose narrative;
- some sections of a text may be written in one form, and others in a different form: for example, a novel where some chapters are in prose, others take the form of dictionary entries, and still others the form of scenes in a play;
- an encoded text may include detailed analytic annotation, for example of rhetorical or linguistic features;
- an encoded text may combine a literal transcription with a diplomatic edition of the same or different sources;
- the description of a text may require additional specialized metadata elements, for example when describing manuscript material in detail.

The TEI provides mechanisms to support all of these and many other use cases. The architecture permits elements and attributes from any combination of modules to co-exist within a single schema. Within particular modules, elements and attributes are provided to support differing views of the 'granularity' of a text, for example:

- a definition of a corpus or collection as a series of <TEI> documents, sharing a common TEI header (see chapter *16. Language Corpora*)
- a definition of composite texts which combine optional front- and back-matter with a group of collected texts, themselves possibly composite (see section *4.3.1. Grouped Texts*)
- an element for the representation of *embedded texts*, where one narrative appears to 'float' within another (see section *4.3.2. Floating Texts*)

Subsequent chapters of these Guidelines describe in detail markup constructs appropriate for these and many other possible features of interest. The markup constructs can be combined as needed for any given set of applications or project.

For example, a project aiming to produce an ambitious digital edition of a collection of manuscript materials, to include detailed metadata about each source, digital images of the content, along with a detailed transcription of each source, and a supporting biographical and geographical database might need a schema combining several modules, as follows:

```
<schemaSpec ident="TEI-PROJECT" start="TEI">
  <moduleRef key="tei"/>
  <moduleRef key="header"/>
  <moduleRef key="core"/>
  <moduleRef key="textstructure"/>
  <moduleRef key="msdescription"/>
<!-- manuscript description -->
  <moduleRef key="transcr"/>
<!-- transcription of primary sources -->
  <moduleRef key="figures"/>
<!-- figures and tables -->
  <moduleRef key="namesdates"/>
<!-- names, dates, people, and places -->
</schemaSpec>
```

Alternatively, a simpler schema might be used for a part of such a project: those preparing the transcriptions, for example, might need only elements from the `core`, `textstructure`, and `transcr` modules, and might therefore prefer to use a simpler schema such as that generated by the following:

```
<schemaSpec ident="TEI-TRANSCR" start="TEI">
  <moduleRef key="tei"/>
  <moduleRef key="core"/>
  <moduleRef key="textstructure"/>
  <moduleRef key="transcr"/>
</schemaSpec>
```

The TEI architecture also supports more detailed customization beyond the simple selection of modules. A schema may suppress elements from a module, suppress some of their attributes, change their names, or even add new elements and attributes. Detailed discussion of the kind of modification possible in this way is provided in *24.3. Customization* and conformance rules relating to their application are discussed in *24.4. Conformance*. These facilities are available for any schema language (though some features may not be available in all languages). The ODD language also makes it possible to combine TEI and non-TEI modules into a single schema, provided that the non-TEI module is expressed using the RELAX NG schema language (see further *23.8.2. Combining TEI and Non-TEI Modules*).

## 1.3 The TEI Class System

The TEI scheme distinguishes about five hundred different elements. To aid comprehension, modularity, and modification, the majority of these elements are formally classified in some way. Classes are used to express two distinct kinds of commonality among elements. The elements of a class may share some set of attributes, or they may appear in the same locations in a content model. A class is known as an *attribute class* if its members share attributes, and as a *model class* if its members appear in the same locations. In either case, an element is said to *inherit* properties from any classes of which it is a member.

Classes (and therefore elements which are members of those classes) may also inherit properties from other classes. For example, supposing that class A is a member (or a *subclass*) of class B, any element which is a member of class A will inherit not only the properties defined by class A, but also those defined by class B. In such a situation, we also say that class B is a *superclass* of class A. The properties of a superclass are inherited by all members of its subclasses.

A basic understanding of the classes into which the TEI scheme is organized is strongly recommended and is essential for any successful customization of the system.

### 1.3.1 Attribute Classes

An attribute class groups together elements which share some set of common attributes. Attribute classes are given names composed of the prefix `att.`, often followed by an adjective. For example, the members of the class `att.canonical` have in common a *key* and a *ref* attribute, both of which are inherited from their membership in

the class rather than individually defined for each element. These attributes are said to be defined by (or inherited from) the `att.canonical` class. If another element were to be added to the TEI scheme for which these attributes were considered useful, the simplest way to provide them would be to make the new element a member of the `att.canonical` class. Note also that this method ensures that the attributes in question are always defined in the same way, taking the same default values etc., no matter which element they are attached to.

Some attribute classes are defined within the `tei` infrastructural module and are thus globally available. Other attribute classes are specific to particular modules and thus defined in other chapters. Attributes defined by such classes will not be available unless the module concerned is included in a schema.

The attributes provided by an attribute class are those specified by the class itself, either directly, or by inheritance from another class. For example, the attribute class `att.pointing.group` provides attributes *domains* and *targFunc* to all of its members. This class is however a subclass of the `att.pointing` class, from which its members also inherit the attributes *target*, *targetLang* and *evaluate*. Members of the class `att.pointing` will thus have these three attributes, while members of the class `att.pointing.group` will have all five.

Note that some modules define superclasses of an existing infrastructural class. For example, the global attribute class `att.divLike` makes attributes *org* and *sample* available, while the `att.metrical` class, which is specific to the `verse` module, provides attributes *met*, *real*, and *rhyme*. Because `att.metrical` is defined as a superclass of `att.divLike`, all five of these attributes are available to elements; the declaration for `att.metrical` adds its three attributes to the three already defined by `att.divLike` when the `verse` module is included in a schema. If, however, this module is not included in a schema, then the `att.divLike` class supplies only the two attributes first mentioned.

Attributes specific to particular modules are documented along with the relevant module rather than in the present chapter. One particular attribute class, known as `att.global`, is common to all modules, and is therefore described in some detail in the next section. A full list of all attribute classes is given in *Appendix B Attribute Classes* below.

#### 1.3.1.1 Global Attributes

The following attributes are defined in the infrastructure module for every TEI element.

`att.global` provides attributes common to all elements in the TEI encoding scheme.

- `@xml:id` (identifier) provides a unique identifier for the element bearing the attribute.
- `@n` (number) gives a number (or other label) for an element, which is not necessarily unique within the document.
- `@xml:lang` (language) indicates the language of the element content using a 'tag' generated according to BCP 47.
- `@rend [att.global.rendition]` (rendition) indicates how the element in question was rendered or presented in the source text.
- `@style [att.global.rendition]` contains an expression in some formal style definition language which defines the rendering or presentation used for this element in the source text.
- `@rendition [att.global.rendition]` points to a description of the rendering or presentation used for this element in the source text.
- `@xml:base` provides a base URI reference with which applications can resolve relative URI references into absolute URI references.
- `@xml:space` signals an intention about how white space should be managed by applications.
- `@source [att.global.source]` specifies the source from which some aspect of this element is drawn.
- `@cert [att.global.responsibility]` (certainty) signifies the degree of certainty associated with the intervention or interpretation.
- `@resp [att.global.responsibility]` (responsible party) indicates the agency responsible for the intervention or interpretation, for example an editor or transcriber.

Some of these attributes (specifically *xml:id*, *n*, *xml:lang*, *xml:base* and *xml:space*) are provided by the `att.global` attribute class itself. The others are provided by one its subclasses `att.global.rendition`, `att.global.responsibility`, or `att.global.source`. Their usage is discussed in the following subsections.

Several other globally-available attributes are defined by other subclasses of the `att.global` class. These are

provided by other modules, and are therefore discussed in the chapter discussing that module. A brief summary table is provided in section *1.3.1.1.7. Other Globally Available Attributes* below.

**1.3.1.1.1 Element Identifiers and Labels** The value supplied for the *xml:id* attribute must be a legal *name*, as defined in the World Wide Web Consortium's XML Recommendation. This means that it must begin with a letter, or the underscore character ('_'), and contain no characters other than letters, digits, hyphens, underscores, full stops, and certain combining and extension characters.[34]

In XML names (and thus the values of *xml:id* in an XML TEI document) uppercase and lowercase letters are distinguished, and thus *partTime* and *parttime* are two distinctly different names, and could (though perhaps unwisely) be used to denote two different element occurrences.

If two elements are given the same identifier, a validating XML parser will signal a syntax error. The following example, therefore, is *not* valid:

```
<p xml:id="PAGE1"><q>What's it going to be then, eh?</q></p>
<p xml:id="PAGE1">There was me, that is Alex, and my three droogs, that is Pete,
Georgie, and Dim, ... </p>
```

For a discussion of methods of providing unique identifiers for elements, see section *3.11.2. Creating New Reference Systems*.

The *n* attribute also provides an identifying name or number for an element, but in this case the information need not be a legal *xml:id* value. Its value may be any string of characters; typically it is a number or other similar enumerator or label. For example, the numbers given to the items of a numbered list may be recorded with the *n* attribute; this would make it possible to record errors in the numeration of the original, as in this list of chapters, transcribed from a faulty original in which the number 10 is used twice, and 11 is omitted:

```
<list rend="numbered">
  <item n="1">About These Guidelines</item>
  <item n="2">A Gentle Introduction to XML</item>
  <item n="9">Verse</item>
  <item n="10">Drama</item>
  <item n="10">Spoken Materials </item>
  <item n="12">Dictionaries</item>
</list>
```

The *n* attribute may also be used to record non-unique names associated with elements in a text, possibly together with a unique identifier as in the following example:

```
<div type="book" n="one" xml:id="TXT0101">
<!-- ... -->
  <div type="stanza" n="xlii">
<!-- ... -->
  </div>
</div>
```

As noted above there is no requirement to record a value for either the *xml:id* or the *n* attribute. Any XML processor can identify the sequential position of one element within another in an XML document without any additional tagging. An encoding in which each line of a long poem is explicitly labelled with its numerical sequence such as the following

[34] The colon is also by default a valid name character; however, it has a specific purpose in XML (to indicate namespace prefixes), and may not therefore be used in any other way within a name.

```
<l  n="1">
<!-- ... -->
</l>
<l  n="2">
<!-- ... -->
</l>
<l  n="3">
<!-- ... -->
</l>
<!-- ... -->
<l  n="100">
<!-- ... -->
</l>
```

is therefore probably redundant.

**1.3.1.1.2 Language Indicators** The *xml:lang* attribute indicates the natural language and writing system applicable to the content of a given element. If it is not specified, the value is inherited from that of the immediately enclosing element. As a rule, therefore, it is simplest to specify the base language of the text on the <TEI> element, and allow most elements to take the default value for *xml:lang*; the language of an element then need be explicitly specified only for elements in languages other than the base language. For this reason, it is recommended practice to supply a default value for the *xml:lang* attribute, either on the <TEI> root element, or on both the <teiHeader> and the <text> element. The latter is appropriate in the not uncommon case where the text element in a TEI document uses a different default language from that of the TEI header attached to it. Other language shifts in the source should be explicitly identified by use of the *xml:lang* attribute on an element at an appropriate level wherever possible.

In the following example schematic, an English language TEI header is attached to an English language text:

```
<TEI  xml:lang="en" xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!-- ... -->
  </teiHeader>
  <text>
<!-- ... -->
  </text>
</TEI>
```

The same effect would be obtained by specifying the default language for both header and text:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader  xml:lang="en">
<!-- ... -->
  </teiHeader>
  <text  xml:lang="en">
<!-- ... -->
  </text>
</TEI>
```

The latter approach is necessary in the case where the two differ: for example, where an English language header is applied to a French text:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader  xml:lang="en">
```

```
<!-- ... -->
  </teiHeader>
  <text  xml:lang="fr">
<!-- ... -->
  </text>
</TEI>
```

The same principle applies at any hierarchic level. In the following example, the default language of the text is French, but one section of it is in German:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader  xml:lang="en">
<!-- ... -->
  </teiHeader>
  <text  xml:lang="fr">
    <body>
      <div>
<!-- chapter one is in French -->
      </div>
      <div  xml:lang="de">
<!-- chapter two is in German -->
      </div>
      <div>
<!-- chapter three is French -->
      </div>
<!-- ... -->
    </body>
  </text>
</TEI>
```

Similarly, in the following example the *xml:lang* attribute on the <term> element allows us to record the fact that the technical terms used are Latin rather than English; no *xml:lang* attribute is needed on the <q> element, by contrast, because it is in the same language as its parent.

```
<p  xml:lang="en">The
constitution declares <q>that no bill of attainder or <term  xml:lang="la">ex post
        facto</term> law shall be passed.</q> ...</p>
```

Note that in cases where it is advisable or necessary to identify the language of the text that is pointed at, the (non-global) attribute *targetLang* should be used, for example in

```
<ptr  target="x12"  targetLang="fr"/>
```

the pointer references text written in French.

The values used for the *xml:lang* and *targetLang* attributes must be constructed in a particular way, using values from standard lists. See further *vi.1 Language Identification*.

Additional information about a particular language may be supplied in the <language> element within the header (see section *2.4.2. Language Usage*).

**1.3.1.1.3 Rendition Indicators** The *rend*, *rendition*, and *style* attributes are all used to give information about the physical presentation of the text in the source. In the following example, *rend* is used to indicate that both the emphasized word and the proper name are printed in italics:

```
<p> ... Their motives <emph  rend="italics">might</emph> be pure
and pious; but he was equally alarmed by his knowledge of the ambitious
<name  rend="italics">Bohemond</name>, and his ignorance of the Transalpine chiefs:
...</p>
```

The same effect might be achieved using the *style* attribute, as follows:

```
<p> ... Their motives <emph  style="font-style: italic">might</emph> be pure and pious;
but he was equally alarmed by his knowledge of
the ambitious <name  style="font-style: italic">Bohemond</name>, and his ignorance of
the Transalpine chiefs: ...</p>
```

If all or most <emph> and <name> elements are rendered in the text by italics, it will be more convenient to register that fact in the TEI header once and for all (using the <rendition> element discussed below) and specify a *rend* or *style* value only for any elements which deviate from the stated rendition.

The main difference between *rend* attribute and *style* is that the value used for the former may contain one or more tokens from any vocabulary devised by the encoder, separated by space characters, whereas the value used for the latter must be a single string taken from a formally-defined style definition language such as CSS. The *rend* attribute values are sequence-indeterminate set of whitespace-separated tokens, whereas *style* values allow whitespace and sequence relationships as part of the formally-defined style definition language.

The <rendition> element defined in *2.3.4. The Tagging Declaration* may be used to hold repeatedly-used format descriptions. A <rendition> element can then be associated with any element, either by default, or by means of the global *rendition* attribute. For example:

```
<tagsDecl>
<!-- define italic style using CSS, selecting it as default for emph and hi elements
-->
  <rendition  xml:id="IT"  scheme="css"
    selector="emph, hi">font-style: italic;</rendition>
<!-- define a serif font family, selecting it as default for the text element -->
  <rendition  xml:id="FontRoman"  scheme="css"
    selector="text">font-family: serif;</rendition>
</tagsDecl>
<!-- ... -->
<text>
  <body>
    <div>
      <p  rendition="#IT">
<!-- this paragraph uses the seriffed font, but is in italic-->
      </p>
      <p>
<!-- this paragraph uses the seriffed font, but is not in italic -->
      </p>
    </div>
  </body>
</text>
```

The *rendition* attribute always points to one or more <rendition> elements, each of which defines some aspect of the rendering or appearance of the text in its original form. These details may most conveniently be described using a formal style definition language, such as CSS (?) or XSL-FO (?); in some other formal language developed for a specific project; or even informally in running prose. Although languages such as CSS and XSL-FO are generally used to describe document output to screen or print, they nonetheless provide formal and precise mechanisms for describing the appearance of source documents, especially print documents, but also many aspects of manuscript

documents. For example, both CSS and XSL-FO provide mechanisms for describing typefaces, weight, and styles; character and line spacing; and so on.

As noted above, the *style* attribute is provided for encoders wishing to describe the appearance of individual source elements using a language such as CSS directly rather than by reference to a <rendition> element. Its value may be any expression in the chosen formal style definition language.

Formal definition languages such as CSS typically identity a series of *properties* (such as font-style or margin-left) for which *values* are specified. A sequence of such property-value pairs makes up a stylesheet. The TEI uses such languages simply to describe the appearance of a source document, rather than to control how it should be formatted.

In the TEI scheme, it is possible to supply information about the appearance of elements within a source document in the following distinct ways:

1. One or more properties may be specified as the default for a set of elements (based on an external scheme, by default CSS), using <rendition> elements and their *selector* attributes;
2. One or more properties may be specified for individual element occurrences, using the *rend* attribute with any convenient set of one or more sequence-indeterminate tokens;
3. One or more properties may be specified for individual element occurrences, using the *rendition* attribute to point to <rendition> elements;
4. One or more properties may be supplied explicitly for individual element occurrences, using the *style* attribute.

If the same property is specified in more than one of the above ways, the one with the highest number in the list above is understood to be applicable. The resulting properties from each way are then combined to provide the full set of property-value pairs applicable to the given element, and (by default) to all of its children.

For simplicity of processing, the same formal style definition should be used throughout; however, the architecture does permit this to be varied, by using the *scheme* attribute to indicate a different language for one or more <rendition> elements. Care should be taken to ensure that such values can be meaningfully combined. Similar considerations apply to the use of the *rend* attribute, if this is used in combination with either *rendition* or *style*.

Note that these TEI attributes always describe the rendition or appearance of the source document, *not* intended output renditions, although often the two may be closely related.

**1.3.1.1.4 Sources, certainty, and responsibility** The *source* attribute is used to indicate the source of an element and its content, for example by pointing to a bibliographic citation for a quotation to indicate the source from which it derives. The target of the pointer may be an entry in a bibliographic list of some kind, or a pointer to a digital version of the source itself.

As with other TEI pointers, the value of this attribute is expressed as any form of URI, for example an absolute URL, a relative URL, or a private scheme URI that is expanded to a relative or absolute URL as documented in a <prefixDef>. In the following typical example a relative 'bare name' URL value is used to point to a <bibl> elsewhere in the bibliography of the document which contains a bibliographic source for the quotation itself:

```
<p>
<!-- ... -->
  <quote source="#chicago-15_ed">Grammatical theories
     are in flux, and the more we learn, the less we
     seem to know.</quote>
<!-- ... -->
</p>
<!-- ... -->
<bibl xml:id="chicago-15_ed">
  <title level="m">The Chicago Manual of Style</title>,
<edition>15th edition</edition>.
<pubPlace>Chicago</pubPlace>:
<publisher>University of Chicago Press</publisher>
(<date>2003</date>),
<biblScope unit="page">p.147</biblScope>.

</bibl>
```

Alternatively, the quotation might be directly linked to the online edition of this source using a full URI :

```
<p>
<!-- ... -->
  <quote  source="http://www.chicagomanualofstyle.org/15/ch05/ch05_sec002.html">Grammatical
theories
    are in flux, and the more we learn, the less we
    seem to know.</quote>
<!-- ... -->
</p>
```

The *source* attribute is also used on schema documentation elements such as <schemaSpec> or <elementRef> to indicate the location from which declarations for the components being defined may be obtained by an ODD processor. For example, a customization wishing to include the <p> element specifically as it was in version 2.0.1 of TEI P5 would indicate the source for this on an <elementRef> element like the following:

```
<elementRef  key="p"  source="tei:2.0.1"/>
```

Here the value of the *source* attribute is provided using private URI syntax, using a short cut predefined for the TEI Guidelines. More generally, an ODD customization can point to a URI from which a compiled version of any ODD can be downloaded. The above shortcut is equivalent to

```
<elementRef  key="p"
  source="http://www.tei-c.org/Vault/P5/2.0.1/xml/tei/odd/p5subset.xml"/>
```

Elements such as <moduleRef> or <elementRef> can use the *source* attribute in this way to point to any previously compiled set of TEI ODD specifications which are to be included in a schema, as further discussed in section *23.8.1. TEI customizations*

The *cert* attribute provides a method of indicating the encoder's certainty concerning an intervention or interpretation represented by the markup. It is typically used where the encoder wishes to supply one or more possible corrections to a text, indicating the certainty they wish to attach to each, as in the following example:

```
 Blessed are the <choice>
  <sic>cheesemakers</sic>
  <corr  cert="high">peacemakers</corr>
  <corr  cert="low">placemakers</corr>
</choice>:
for they shall be called the children of God.
```

The *cert* attribute will usually, as here, characterize the degree of certainty simply as high, medium or low. In situations where a more detailed or nuanced indication is required, it can instead supply a probability value between 0 (minimal probability) and 1 (maximal probability). Other more sophisticated mechanisms are discussed in chapter *22. Certainty, Precision, and Responsibility*.

The *resp* attribute is used to indicate the person or organization considered responsible for some aspects of the information encoded by an element. For example, the preceding example might be revised as follows to indicate the editors responsible for the two corrections:

```
 ... <corr  cert="high"  resp="#ed1">peacemakers</corr>
<corr  cert="low"  resp="#ed2">placemakers</corr>...
```

When a more detailed or nuanced representation of responsibility is required, it is recommended that the element indicated by the *resp* attribute should not be a generic agent (for example a <person> or <org>) but a more precise element such as <respStmt>, <author>, or <editor> which can document the exact role played by the agent. In the following example, we indicate that the correction of n to u was made by a particular named transcriber:

```
<!-- in the <text> ... --><lg>
<!-- ... -->
  <l>Punkes, Panders, baſe extortionizing
     sla<choice>
      <sic>n</sic>
      <corr  resp="#JENSJ">u</corr>
    </choice>es,</l>
<!-- ... -->
</lg>
<!-- in the <teiHeader> ... -->
<!-- ... -->
<respStmt  xml:id="JENSJ">
  <resp>Transcriber</resp>
  <name>Janelle Jenstad</name>
</respStmt>
```

Pointing to multiple <respStmt>s allows the encoder to specify clearly each of the roles played in part of a TEI file (creating, transcribing, encoding, editing, proofing etc.). If appropriate, the <name> element inside a <respStmt> may also be associated with a more detailed <person> or <org> element using methods discussed in chapter *14. Names, Dates, People, and Places*.

**1.3.1.1.5 Evaluation of Links** Several TEI elements carry attributes whose values are defined as `anyURI`, meaning that such attributes supply a link or pointer, typically expressed as a URL. Like other XML applications, the TEI allows use of a special attribute to set the context within which relative URLs are to be evaluated. The global attribute *xml:base* is defined as part of the XML specification and belongs to the XML namespace rather than the TEI namespace. We do not describe it in detail here: reference information about *xml:base* is provided by **?**

In essence *xml:base* is used to set a context for all relative URLs within the scope of the element on which it is specified. For example:

```
<body>
  <div  xml:base="http://www.example.org/somewhere.xml">
    <p>
<!--... -->
     <ptr  target="elsewhere.xml"/>
<!--... -->
    </p>
  </div>
  <div>
    <p>
<!--... -->
     <ptr  target="elsewhere.xml"/>
<!--... -->
    </p>
  </div>
</body>
```

The first <ptr> element here is within the scope of a  which supplies a value for *xml:base*; its target is therefore to be found at `http://www.example.org/elsewhere.xml`. The second <ptr>, however, is within the scope of a  which does not change the default context, and its target is therefore a document in the same directory as the current document.

The *xml:base* attribute is intended to enable the stable resolution of relative URIs in a document after that document's context may have changed (for example as a result of being embedded in another document via XInclude). Setting the *xml:base* simply as a way to allow encoders to write shorter URIs is not recommended. In particular, *xml:base* may cause ambiguity as to the referent of same-document references in the form `#id` (where `id` is an *xml:id*). RFC 3986 states that URIs of this type should not result in the loading of a different document. The RFC therefore assumes that such references are internal to the document in which they are located. Using *xml:base* to denote arbitrary external bases while also using same-document references may mean that software agents deal with these links in unexpected and inconsistent ways. Further discussion of this attribute and its effect on TEI linking methods is provided in chapter *17. Linking, Segmentation, and Alignment*.

**1.3.1.1.6 XML Whitespace** The global attribute *xml:space* provides a mechanism for indicating to systems processing an XML file how they should treat whitespace, that is, any sequences of consecutive tab (#x09), space (#x20), carriage return (#x0D) or linefeed (#x0A) characters. Like *xml:id* this attribute is defined as part of the XML specification and belongs to the XML namespace rather than the TEI namespace. Complete information about this attribute is provided by section 2.10 of the XML Specification; here we provide a summary of how its use affects users of the TEI scheme.

The *xml:space* attribute has only two permitted values: preserve and default. The first indicates that whitespace in a text node—every carriage return, every tab, etc.—should be maintained as is when the document is processed. The second (which is implied when the attribute is not supplied), indicates that whitespace should be handled 'as appropriate'. Exactly what is deemed appropriate is left unspecified by the XML Recommendation.

These Guidelines assume one of two different ways of processing whitespace will apply in a given case, depending on an element's content model. For an element that can contain only other elements with no intervening non-whitespace characters, whitespace is considered to have no semantic significance, and should therefore be discarded by a processor. For example, in a <choice> element, such as

```
<choice>
  <sic>1724</sic>
  <corr>1728</corr>
</choice>
```

since non-whitespace text is not permitted between the <choice> start-tag and the <sic> tags or between the <sic> and <corr> tags, any whitespace found there has no significance and can be ignored completely by a processor.

Similarly, the <address> element has a content model containing only elements: any punctuation or whitespace required between the lines of an address must therefore be supplied by the processor, as any whitespace present in the input document will be ignored.

Elements with content models of this type are comparatively unusual in the TEI: a list of them is provided in the TEI release file `stripspace.xsl.model`, formatted there for use as an `<xsl:strip-space>` command for XSL stylesheets.

Most TEI elements permit what is known as mixed-content: that is, they can contain both text and other elements. Here the assumption of these Guidelines is that whitespace will be normalized. This means that all space, carriage return, linefeed, and tab characters are converted into spaces, all consecutive spaces are then deleted and replaced by one space, and then space immediately after a start-tag or immediately before an end-tag is deleted. The result is that this encoding,

```
<persName>
  <forename>Edward</forename>
  <forename>George</forename>
  <surname  type="linked">Bulwer-Lytton</surname>, <roleName>Baron Lytton of
  <placeName>Knebworth</placeName>
  </roleName>
</persName>
```

would be rendered as 'Edward George Bulwer-Lytton, Baron Lytton of Knebworth'. The space before his name has been removed, a space is included between his forenames, the comma is preserved, and the newlines within his name have all been removed.

If the default treatment described above is not appropriate for a mixed content element, the processing required may be described in the <encodingDesc> element of the TEI header, but generic XML processing tools may not take note of this.

Alternatively, the *xml:space* attribute may be supplied with a value of preserve in order to indicate that every space, tab, carriage return and linefeed character found within that element in the document being processed is significant. Typically, the result of that processing will be to retain the whitespace characters in the output. Thus if the above example began <persName xml:space="preserve">, the resulting text would most likely be rendered over five lines, indented, and with a blank line following.

The `xml:space="preserve"` attribute is rarely used in TEI documents because such layout features are generally captured with less risk and more precision by using native TEI elements such as <lb> or <space>, or by using the renditional attributes described in section *1.3.1.1.3. Rendition Indicators*.

**1.3.1.1.7 Other Globally Available Attributes** The following table lists for convenience other potentially available global attributes. The table specifies the name of the attribute class providing the attributes concerned, the module which must be included in a schema if the attributes are to be made available, and the section of these Guidelines where the class is discussed.

| class name | module name | see further |
|---|---|---|
| att.global.linking | linking | *17. Linking, Segmentation, and Alignment* |
| att.global.analytic | analysis | *18. Simple Analytic Mechanisms* |
| att.global.facs | transcr | *12.1. Digital Facsimiles* |
| att.global.change | transcr | *12.7. Identifying Changes and Revisions* |

## 1.3.2 Model Classes

As noted above, the members of a given TEI model class share the property that they can all appear in the same location within a document. Wherever possible, the content model of a TEI element is expressed not directly in terms of specific elements, but indirectly in terms of particular model classes. This makes content models simpler and more consistent; it also makes them much easier to understand and to modify.

Like attribute classes, model classes may have subclasses or superclasses. Just as elements inherit from a class the ability to appear in certain locations of a document (wherever the class can appear), so all members of a subclass inherit the ability to appear wherever any superclass can appear. To some extent, the class system thus provides a way of reducing the whole TEI galaxy of elements into a tidy hierarchy. This is however not entirely the case.

In fact, the nature of a given class of elements can be considered along two dimensions: as noted, it defines a set of places where the class members are permitted within the document hierarchy; it also implies a semantic grouping of some kind. For example, the very large class of elements which can appear within a paragraph comprises a number of other classes, all of which have the same structural property, but which differ in their field of application. Some are related to highlighting, while others relate to names or places, and so on. In some cases, the 'set of places where class members are permitted' is very constrained: it may just be within one specific element, or one class of element, for example. In other cases, elements may be permitted to appear in very many places, or in more than one such set of places.

These factors are reflected in the way that model classes are named. If a model class has a name containing part, such as `model.divPart` or `model.biblPart` then it is primarily defined in terms of its structural location. For example, those elements (or classes of element) which appear as content of a  constitute the `model.divPart` class; those which appear as content of a <bibl> constitute the `model.biblPart` class. If, however, a model class has a name containing like, such as `model.biblLike` or `model.nameLike`, the implication is that its members all have some additional semantic property in common, for example containing a bibliographic description, or containing some form of name, respectively. These semantically-motivated classes often provide a useful way of dividing up large structurally-motivated classes: for example, the very general structural class `model.pPart.data` ('data elements that form part of a paragraph') has four semantically-motivated member classes (`model.addressLike`, `model.dateLike`, `model.measureLike`, and `model.nameLike`), the last of these being itself a superclass with several members.

Although most classes are defined by the `tei` infrastructure module, a class cannot be populated unless some other specific module is included in a schema, since element declarations are contained by modules. Classes are not

declared 'top down', but instead gain their members as a consequence of individual elements' declaration of their membership. The same class may therefore contain different members, depending on which modules are active. Consequently, the content model of a given element (being expressed in terms of model classes) may differ depending on which modules are active.

Some classes contain only a single member, even when all modules are loaded. One reason for declaring such a class is to make it easier for a customization to add new member elements in a specific place, particularly in areas where the TEI does not make fully elaborated proposals. For example, the TEI class `model.rdgLike`, initially empty, is expanded by the `textcrit` module to include just the TEI <rdg> element. A project wishing to add an alternative way of structuring text-critical information could do so by defining their own elements and adding it to this class.

Another reason for declaring single-member classes is where the class members are not needed in all documents, but appear in the same place as elements which are very frequently required. For example, the specialized element <g> used to represent a non-Unicode character or glyph is provided as the only member of the `model.gLike` class when the `gaiji` module is added to a schema. References to this class are included in almost every content model, since if it is used at all the <g> must be available wherever text is available; however these references have no effect unless the gaiji module is loaded.

At the other end of the scale, a few of the classes predefined by the tei module are subsequently populated with very many members. For example, the class `model.pPart.edit` groups all the classes of element for simple editorial correction and transcription which can appear within a <p> or paragraph element. The `core` module alone adds more than fifty elements to this class; the `namesdates` module adds another twenty, as does the `tagdocs` module. Since the <p> element is one of the basic building blocks of a TEI document it is not surprising that each module will need to add elements to it. The class system here provides a very convenient way of controlling the resulting complexity. Typically, elements are not added directly to these very general classes, but via some intermediate semantically-motivated class.

Just as there are a few classes which have a single member, so there are some classes which are used only once in the TEI architecture. These classes, which have no superclass and therefore do not fit into the class hierarchy defined here, are a convenient way of maintaining elements which are highly structured internally, but which appear from the outside to be uniform objects like others at the same level.[35] Members of such classes can only ever appear within one element, or one class of elements. For example, the class `model.addrPart` is used only to express the content model for the element <address>; it references some other classes of elements, which can appear elsewhere, and also some elements which can only appear inside an address.

#### 1.3.2.1 Informal Element Classifications

Most TEI elements may also be informally classified as belonging to one of the following groupings:

***divisions*** high level, possibly self-nesting, major divisions of texts. These elements populate such classes as `model.divLike` or `model.div1Like`, and typically form the largest component units of a text.

***chunks*** elements such as paragraphs and other paragraph-level elements, which can appear directly within texts or within divisions of them, but not (usually) within other chunks. These elements populate the class `model.divPart`, either directly or by means of other classes such as `model.pLike` (paragraph-like elements), `model.entryLike`, etc.

***phrase-level elements*** elements such as highlighted phrases, book titles, or editorial corrections which can occur only within chunks, but not between them (and thus cannot appear directly within a division). These elements populate the class `model.phrase`.[36]

The TEI also identifies two further groupings derived from these three:

***inter-level elements*** elements such as lists, notes, quotations, etc. which can appear either between chunks (as children of a ) or within them; these elements populate the class `model.inter`. Note that this class is not a superset of the `model.phrase` and `model.divPart` classes but rather a distinct grouping of elements which are both chunk-like and phrase-like. However, the classes `model.phrase`, `model.pLike`, and `model.inter` are all disjoint.

[35] In former editions of these Guidelines, such elements were known metaphorically as 'crystals'.

[36] Note that in this context, *phrase* means any string of characters, and can apply to individual words, parts of words, and groups of words indifferently; it does not refer only to linguistically-motivated phrasal units. This may cause confusion for readers accustomed to applying the word in a more restrictive sense.

| Content model | Number of elements using this | Description |
|---|---|---|
| macro.phraseSeq | 87 | defines a sequence of character data and phrase-level elements. |
| macro.paraContent | 52 | defines the content of paragraphs and similar elements. |
| macro.specialPara | 34 | defines the content model of elements such as notes or list items, which either contain a series of component-level elements or else have the same structure as a paragraph, containing a series of phrase-level and inter-level elements. |
| macro.phraseSeq.limited | 25 | defines a sequence of character data and those phrase-level elements that are not typically used for transcribing extant documents. |
| macro.xtext | 9 | defines a sequence of character data and gaiji elements. |
| macro.limitedContent | 7 | defines the content of prose elements that are not used for transcription of extant materials. |
| macro.abContent | 1 | defines the content of anonymous block elements. |

Table 4:

***components*** elements which can appear directly within texts or text divisions; this is a combination of the inter- and chunk- level elements defined above. These elements populate the class `model.common`, which is defined as a superset of the classes `model.divPart`, `model.inter`, and (when the dictionary module is included in a schema) `model.entryLike`.

Broadly speaking, the front, body, and back of a text each comprises a series of components, optionally grouped into divisions.

As noted above, some elements do not belong to any model class, and some model classes are not readily associated with any of the above informal groupings. However, over two-thirds of the 588 elements defined in the present edition of these Guidelines are classified in this way, and future editions of these recommendations will extend and develop this classification scheme.

A complete alphabetical list of all model classes is provided in *Appendix A Model Classes*.

## 1.4 Macros

The infrastructure module defined by this chapter also declares a number of *macros*, or shortcut names for frequently occurring parts of other declarations. Macros are used in two ways in the TEI scheme: to stand for frequently-encountered content models, or parts of content models (*1.4.1. Standard Content Models*); and to stand for attribute datatypes (*1.4.2. Datatype Specifications*).

### 1.4.1 Standard Content Models

As far as possible, the TEI schemas use the following set of frequently-encountered content models to help achieve consistency among different elements.

`macro.paraContent` (paragraph content) defines the content of paragraphs and similar elements.

`macro.limitedContent` (paragraph content) defines the content of prose elements that are not used for transcription of extant materials.

`macro.phraseSeq` (phrase sequence) defines a sequence of character data and phrase-level elements.

`macro.phraseSeq.limited` (limited phrase sequence) defines a sequence of character data and those phrase-level elements that are not typically used for transcribing extant documents.

`macro.specialPara` ('special' paragraph content) defines the content model of elements such as notes or list items, which either contain a series of component-level elements or else have the same structure as a paragraph, containing a series of phrase-level and inter-level elements.

`macro.xtext` (extended text) defines a sequence of character data and gaiji elements.

The present version of the TEI Guidelines includes some 588 different elements. Table 1 shows, in descending order of frequency, the seven most commonly used content models.

### 1.4.2 Datatype Specifications

The values which attributes may take in a TEI schema are defined, for the most part, by reference to a TEI *datatype specification*. Each such specification is defined in terms of other primitive datatypes, derived mostly from W3C Schema Datatypes, literal values, or other datatypes. This indirection makes it possible for a TEI application to set constraints either globally or in individual cases, by redefining the datatype definition or the reference to it respectively. In some cases, the TEI datatype includes additional usage constraints which cannot be enforced by existing schema languages, although a TEI-compliant processor should attempt to validate them (see further discussion in chapter *24.4. Conformance*).

The following element is used to define a TEI datatype:

`<dataSpec>` (datatype specification) documents a datatype.

TEI-defined datatypes may be grouped into those which define normalized values for numeric quantities, probabilities, or temporal expressions, those which define various kinds of shorthand codes or keys, and those which define pointers or links.

The following datatypes are used for attributes which are intended to hold normalized values of various kinds. First, expressions of quantity or probability:

**teidata.certainty** defines the range of attribute values expressing a degree of certainty.

**teidata.probability** defines the range of attribute values expressing a probability.

**teidata.numeric** defines the range of attribute values used for numeric values.

**teidata.interval** defines attribute values used to express an interval value.

**teidata.count** defines the range of attribute values used for a non-negative integer value used as a count.

Examples of attributes using the `teidata.probability` datatype include *degree* on <damage> or <certainty>; examples of `teidata.numeric` include *quantity* on members of the `att.measurement` class or *value* on <numeric>; examples of `teidata.count` include *cols* on <cell> and <table>.

Next, the datatypes used for attributes which are intended to hold normalized dates or times, durations, truth values, and language identifiers:

**teidata.duration.w3c** defines the range of attribute values available for representation of a duration in time using W3C datatypes.

**teidata.temporal.w3c** defines the range of attribute values expressing a temporal expression such as a date, a time, or a combination of them, that conform to the W3C XML Schema Part 2: Datatypes Second Edition specification.

**teidata.temporal.working** defines the range of values, conforming to the W3C XML Schema Part 2: Datatypes Second Edition specification, expressing a date or a date and a time within the working life of the document.

**teidata.truthValue** defines the range of attribute values used to express a truth value.

**teidata.xTruthValue** (extended truth value) defines the range of attribute values used to express a truth value which may be unknown.

**teidata.language** defines the range of attribute values used to identify a particular combination of human language and writing system.

Note that in each of these cases the values used are those recommended by existing international standards: ISO 8601 as profiled by XML Schema Part 2: Datatypes Second Edition in the case of durations, times, and date; W3C Schema datatypes in the case of truth values; and BCP 47 in the case of language.

The following datatypes have more specialized uses:

**teidata.namespace** defines the range of attribute values used to indicate XML namespaces as defined by the W3C Namespaces in XML Technical Recommendation.

**teidata.namespaceOrName** defines attribute values which contain either an absolute namespace URI or a qualified XML name.

**teidata.outputMeasurement** defines a range of values for use in specifying the size of an object that is intended for display.

**teidata.pattern** defines attribute values which are expressed as a regular expression.

**teidata.point** defines the data type used to express a point in cartesian space.

**teidata.pointer** defines the range of attribute values used to provide a single URI, absolute or relative, pointing to some other resource, either within the current document or elsewhere.

**teidata.authority** defines attribute values which derive from an authority list, which may be an enumerated list defined in the document's schema, a list or taxonomy elsewhere in the document, or an online taxonomy, gazetteer, or other authority.

**teidata.version** defines the range of attribute values which may be used to specify a TEI or Unicode version number.

**teidata.versionNumber** defines the range of attribute values used for version numbers.

**teidata.replacement** defines attribute values which contain a replacement template.

**teidata.xpath** defines attribute values which contain an XPath expression.

By far the largest number of TEI attributes take values which are coded values or names of some kind. These values may be constrained or defined in a number of different ways, each of which is given a different name, as follows:

**teidata.word** defines the range of attribute values expressed as a single word or token.

**teidata.text** defines the range of attribute values used to express some kind of identifying string as a single sequence of Unicode characters possibly including whitespace.

**teidata.name** defines the range of attribute values expressed as an XML Name.

**teidata.enumerated** defines the range of attribute values expressed as a single XML name taken from a list of documented possibilities.

**teidata.gender** defines the range of attribute values used to represent the gender of a person, persona, or character.

**teidata.sex** defines the range of attribute values used to identify the sex of an organism.

**teidata.xmlName** defines attribute values which contain an XML name.

**teidata.prefix** defines a range of values that may function as a URI scheme name.

Attributes of type `teidata.word`, such as *age* on <person>, are used to supply an identifier expressed as any kind of single token or word. The TEI places a few constraints on the characters which may be used for this purpose: only Unicode characters classified as letters, digits, punctuation characters, or symbols can appear in an attribute value of this kind. Note in particular that such values cannot include whitespace characters. Legal values include cholmondeley, été, 1234, e_content, or xml:id, but not grand wazoo. Attributes of this kind are sometimes used to associate (by co-reference) elements of different types.

Where identifiers are defined externally, for example as part of a database or file system, the inability to include whitespace or other special characters in a value may be problematic. In other cases, it may also be simply more convenient to supply a short sequence of natural language words including spaces as a single value. For these reasons, we also provide a datatype `teidata.text` which does permit whitespace and indeed any other Unicode character. Legal values include cholmondeley, été, 1234, e-content, xml:id, and grand wazoo. This datatype should be used with care since XML will not normalize whitespace characters within it: for example the values `n="a  b"` (two spaces) and `n="a   b"` (three spaces) would be considered distinct. This case should be distinguished from that of an attribute permitting multiple values, each of which may be separated by whitespace which *will* be normalized (see further *23.5.3.1. Datatypes*).

Attributes of type `teidata.name` are similar to those of type `teidata.word`, but with the additional constraint that they must be legal XML identifiers, as defined by the XML 1.0 specification, or successors. Hence, they may not begin with digits or punctuation characters. Legal identifiers include cholmondeley, été, e_content, or xml:id, but not grand wazoo or 1234. Attributes of this kind are typically used to represent XML element or attribute names.

Attributes of type `teidata.xmlName` are similar to those of type `teidata.name`, but with the additional constraint that they must not contain a colon character (:, U+003A). Thus attributes of this kind are used to represent XML element or attribute names that do not have a namespace prefix.

Attributes of type `teidata.prefix`, such as *ident* of <prefixDef>, are restricted to strings that form legal URI prefixes.[37] Examples of valid values are http, https, tn3270, xmlrpc.beep, and view-source.

[37] Technically the specification permits the 26 uppercase letters *A-Z*; however, since 'the canonical form is lowercase and documents that specify schemes must do so with lowercase letters', the TEI `teidata.prefix` datatype does not permit uppercase letters.

Attributes of type `teidata.enumerated`, such as *new* on <shift> or *evidence* supplied by `att.editLike`, have the same definition as `teidata.word` above, with the added constraint that the word supplied is taken from a specific list of possibilities. In each case, the element or class specification which includes the definition for the attribute will also contain a list of possible values, together with a prose description of their intended significance. This list may be open (in which case the list is advisory), or closed (in which case it determines the range of legal values). In this latter case, the datatype will not be `teidata.enumerated`, but an explicit list of the possible values.

An attribute may, of course, take more than one value of a given type, for example a list of pointer values, or a list of words. In the TEI scheme, this information is regarded as a property of the <datatype> element used to document the attribute in question rather than as a distinct 'datatype', and is provided by the *minOccurs* or *maxOccurs* attribute. See further *23.5.3.1. Datatypes*.

In a small number of cases, an attribute may take a value of either one datatype or another. These cases are considered as distinct datatypes:

**teidata.probCert** defines a range of attribute values which can be expressed either as a numeric probability or as a coded certainty value.

**teidata.unboundedCount** defines the range of values used for a counting number or the string unbounded for infinity.

**teidata.nullOrName** defines attribute values which contain either the null string or an XML name.

## 1.5 The TEI Infrastructure Module

The `tei` module defined by this chapter is a required component of any TEI schema. It provides declarations for all datatypes, and initial declarations for the attribute classes, model classes, and macros used by other modules in the TEI scheme. Its components are listed below in alphabetical order:

**Module tei: Declarations for classes, datatypes, and macros available to all TEI modules**

- *Classes defined*: att.anchoring att.ascribed att.ascribed.directed att.breaking att.cReferencing att.calendarSystem att.canonical att.citing att.cmc att.damaged att.datable att.datable.w3c att.datcat att.declarable att.declaring att.dimensions att.divLike att.docStatus att.duration.iso att.duration.w3c att.editLike att.edition att.formula att.fragmentable att.global att.global.rendition att.global.responsibility att.global.source att.handFeatures att.indentation att.internetMedia att.interpLike att.locatable att.measurement att.media att.naming att.notated att.partials att.personal att.placement att.pointing att.pointing.group att.ranging att.resourced att.scope att.scoping att.segLike att.sortable att.spanning att.styleDef att.timed att.transcriptional att.typed att.written model.addrPart model.addressLike model.annotationLike model.annotationPart.body model.applicationLike model.attributable model.availabilityPart model.biblLike model.biblPart model.castItemPart model.catDescPart model.certLike model.choicePart model.cmc model.common model.correspActionPart model.correspContextPart model.correspDescPart model.dateLike model.descLike model.describedResource model.dimLike model.div1Like model.div2Like model.div3Like model.div4Like model.div5Like model.div6Like model.div7Like model.divBottom model.divBottomPart model.divGenLike model.divLike model.divPart model.divTop model.divTopPart model.divWrapper model.editorialDeclPart model.egLike model.emphLike model.encodingDescPart model.entryPart model.entryPart.top model.eventLike model.featureVal model.featureVal.complex model.featureVal.single model.frontPart model.frontPart.drama model.gLike model.global model.global.edit model.global.meta model.graphicLike model.headLike model.hiLike model.highlighted model.identEquiv model.identSynonyms model.imprintPart model.inter model.lLike model.lPart model.labelLike model.limitedPhrase model.linePart model.listLike model.measureLike model.milestoneLike model.msItemPart model.msQuoteLike model.nameLike model.nameLike.agent model.noteLike model.objectLike model.oddDecl model.oddRef model.offsetLike model.orgPart model.orgStateLike model.pLike model.pLike.front model.pPart.data model.pPart.edit model.pPart.editorial model.pPart.msdesc model.pPart.transcriptional model.paraPart model.persStateLike model.personLike model.personPart model.phrase model.phrase.xml model.placeLike model.placeNamePart model.placeStateLike model.profileDescPart model.ptrLike model.publicationStmtPart.agency model.publicationStmtPart.detail model.quoteLike model.resource model.respLike model.segLike

model.settingPart model.sourceDescPart model.specDescLike model.stageLike model.standOffPart model.teiHeaderPart model.textDescPart model.titlepagePart

- Macros defined: macro.abContent macro.limitedContent macro.paraContent macro.phraseSeq macro.phraseSeq.limited macro.specialPara macro.xtext teidata.authority teidata.certainty teidata.count teidata.duration.iso teidata.duration.w3c teidata.enumerated teidata.gender teidata.interval teidata.language teidata.name teidata.namespace teidata.namespaceOrName teidata.nullOrName teidata.numeric teidata.outputMeasurement teidata.pattern teidata.point teidata.pointer teidata.prefix teidata.probCert teidata.probability teidata.replacement teidata.sex teidata.temporal.iso teidata.temporal.w3c teidata.temporal.working teidata.text teidata.truthValue teidata.unboundedCount teidata.version teidata.versionNumber teidata.word teidata.xTruthValue teidata.xmlName teidata.xpath

The order in which declarations are made within the infrastructure module is critical, since several class declarations refer to others, which must therefore precede them. Other constraints on the order of declarations derive from the way in which the modularity of the TEI scheme is implemented in different schema languages. The XML DTD fragment implementing this TEI module makes extensive use of *parameter entities* and *marked sections* to effect a kind of conditional construction; the RELAX NG schema fragment similarly predeclares a number of patterns with null ('notAllowed') values. These issues are further discussed in chapter *24.5. Implementation of an ODD System*.

# 2 The TEI Header

This chapter addresses the problems of describing an encoded work so that the text itself, its source, its encoding, and its revisions are all thoroughly documented. Such documentation is equally necessary for scholars using the texts, for software processing them, and for cataloguers in libraries and archives. Together these descriptions and declarations provide an electronic analogue to the title page attached to a printed work. They also constitute an equivalent for the content of the code books or introductory manuals customarily accompanying electronic data sets.

Every TEI-conformant text must carry such a set of descriptions, prefixed to it and encoded as described in this chapter. The set is known as the *TEI header*, tagged <teiHeader>, and has five major parts:

1. a *file description*, tagged <fileDesc>, containing a full bibliographical description of the computer file itself, from which a user of the text could derive a proper bibliographic citation, or which a librarian or archivist could use in creating a catalogue entry recording its presence within a library or archive. The term *computer file* here is to be understood as referring to the whole entity or document described by the header, even when this is stored in several distinct operating system files. The file description also includes information about the source or sources from which the electronic document was derived. The TEI elements used to encode the file description are described in section *2.2. The File Description* below.

2. an *encoding description*, tagged <encodingDesc>, which describes the relationship between an electronic text and its source or sources. It allows for detailed description of whether (or how) the text was normalized during transcription, how the encoder resolved ambiguities in the source, what levels of encoding or analysis were applied, and similar matters. The TEI elements used to encode the encoding description are described in section *2.3. The Encoding Description* below.

3. a *text profile*, tagged <profileDesc>, containing classificatory and contextual information about the text, such as its subject matter, the situation in which it was produced, the individuals described by or participating in producing it, and so forth. Such a text profile is of particular use in highly structured composite texts such as corpora or language collections, where it is often highly desirable to enforce a controlled descriptive vocabulary or to perform retrievals from a body of text in terms of text type or origin. The text profile may however be of use in any form of automatic text processing. The TEI elements used to encode the profile description are described in section *2.4. The Profile Description* below.

4. a container element, tagged <xenoData>, which allows easy inclusion of metadata from non-TEI schemes (i.e., other than elements in the TEI namespace). For example, the MARC record for the encoded document might be included using MARCXML or MODS. A simple set of metadata for harvesting might be included encoded in Dublin Core.

5. a *revision history*, tagged <revisionDesc>, which allows the encoder to provide a history of changes made during the development of the electronic text. The revision history is important for *version control* and for resolving questions about the history of a file. The TEI elements used to encode the revision description are described in section *2.6. The Revision Description* below.

A TEI header can be a very large and complex object, or it may be a very simple one. Some application areas (for example, the construction of language corpora and the transcription of spoken texts) may require more specialized and detailed information than others. The present proposals therefore define both a *core* set of elements (all of which may be used without formality in any TEI header) and some additional elements which become available within the header as the result of including additional specialized modules within the schema. When the module for language corpora (described in chapter *16. Language Corpora*) is in use, for example, several additional elements are available, as further detailed in that chapter.

The next section of the present chapter briefly introduces the overall structure of the header and the kinds of data it may contain. This is followed by a detailed description of all the constituent elements which may be used in the core header. Section *2.7. Minimal and Recommended Headers*, at the end of the present chapter, discusses the recommended content of a minimal TEI header and its relation to standard library cataloguing practices.

## 2.1 Organization of the TEI Header

### 2.1.1 The TEI Header and Its Components

The <teiHeader> element should be clearly distinguished from the *front matter* of the text itself (for which see section *4.5. Front Matter*). A composite text, such as a corpus or collection, may contain several headers, as further discussed below. In the general case, however, a TEI-conformant text will contain a single <teiHeader> element, followed by a single <text> or <facsimile> element, or both.

The header element has the following description:

`<teiHeader>` (TEI header) supplies descriptive and declarative metadata associated with a digital resource or set of resources.

As discussed above, the <teiHeader> element has five principal components:

`<fileDesc>` (file description) contains a full bibliographic description of an electronic file.

`<encodingDesc>` (encoding description) documents the relationship between an electronic text and the source or sources from which it was derived.

`<profileDesc>` (text-profile description) provides a detailed description of non-bibliographic aspects of a text, specifically the languages and sublanguages used, the situation in which it was produced, the participants and their setting.

`<xenoData>` (non-TEI metadata) provides a container element into which metadata in non-TEI formats may be placed.

`<revisionDesc>` (revision description) summarizes the revision history for a file.

Of these, only the <fileDesc> element is required in all TEI headers; the others are optional. That is, only one of the five components of the TEI header (the <fileDesc>) is mandatory, and it also has some mandatory components, as further discussed in *2.2. The File Description* below. The smallest possible valid TEI header thus looks like this:

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>
<!-- title of the resource -->
      </title>
    </titleStmt>
    <publicationStmt>
      <p>
<!-- Information about distribution of the resource -->
      </p>
    </publicationStmt>
    <sourceDesc>
      <p>
<!-- Information about source from which the resource derives -->
      </p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

The content of the elements making up a TEI header may be given in any language, not necessarily that of the text to which the header applies, and not necessarily English. As elsewhere, the *xml:lang* attribute should be used at an appropriate level to specify the language. For example, in the following schematic example, an English text has been given a French header:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader  xml:lang="fr">
<!-- ... -->
  </teiHeader>
  <text  xml:lang="en">
```

```
<!-- ... -->
  </text>
</TEI>
```

In the case of language corpora or collections, it may be desirable to record header information either at the level of the individual components in the corpus or collection, or at the level of the corpus or collection itself (more details concerning the tagging of composite texts are given in section *16. Language Corpora*, which should be read in conjunction with the current chapter). A corpus may thus take the form:

```
<teiCorpus xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!-- corpus-level metadata here -->
  </teiHeader>
  <TEI>
    <teiHeader>
<!-- metadata specific to this text here -->
    </teiHeader>
    <text>
<!-- ... -->
    </text>
  </TEI>
  <TEI>
    <teiHeader>
<!-- metadata specific to this text here -->
    </teiHeader>
    <text>
<!-- ... -->
    </text>
  </TEI>
</teiCorpus>
```

### 2.1.2 Types of Content in the TEI Header

The elements occurring within the TEI header may contain several types of content; the following list indicates how these types of content are described in the following sections:

**free prose** Most elements contain simple running prose at some level. Many elements may contain either prose (possibly organized into paragraphs) or more specific elements, which themselves contain prose. In this chapter's descriptions of element content, the phrase *prose description* should be understood to imply a series of paragraphs, each marked as a <p> element. The word *phrase*, by contrast, should be understood to imply character data, interspersed as need be with phrase-level elements, but not organized into paragraphs. For more information on paragraphs, highlighted phrases, lists, etc., see section *3.1. Paragraphs*.

**grouping elements** Elements whose names end with the suffix *Stmt* (e.g. <editionStmt>, <titleStmt>) and the <xenoData> element enclose a group of specialized elements recording some structured information. In the case of the bibliographic elements, the suffix *Stmt* is used in names of elements corresponding to the 'areas' of the International Standard Bibliographic Description.[38] In the case of the <xenoData> element, the specialized elements are not TEI elements, but rather come from some other metadata scheme. In most cases grouping elements may contain prose descriptions as an alternative to the set of specialized elements, thus allowing the encoder to choose whether or not the information concerned should be presented in a structured form or in prose.

[38]For more information on this highly influential family of standards, first proposed in 1969 by the International Federation of Library Associations, see `http://www.ifla.org/VII/s13/pubs/isbd.htm`. On the relation between the TEI proposals and other standards for bibliographic description, see further section *2.8. Note for Library Cataloguers*.

**declarations** Elements whose names end with the suffix *Decl* (e.g. <tagsDecl>, <refsDecl>) enclose information about specific encoding practices applied in the electronic text; often these practices are described in coded form. Typically, such information takes the form of a series of declarations, identifying a code with some more complex structure or description. A declaration which applies to more than one text or division of a text need not be repeated in the header of each such text or subdivision. Instead, the *decls* attribute of each text (or subdivision of the text) to which the declaration applies may be used to supply a cross-reference to it, as further described in section *16.3. Associating Contextual Information with a Text*.

**descriptions** Elements whose names end with the suffix *Desc* (e.g. <settingDesc>, <projectDesc>) contain a prose description, possibly, but not necessarily, organized under some specific headings by suggested sub-elements.

### 2.1.3 Model Classes in the TEI Header

The TEI header provides a very rich collection of metadata categories, but makes no claim to be exhaustive. It is certainly the case that individual projects may wish to record specialized metadata which either does not fit within one of the predefined categories identified by the TEI header or requires a more specialized element structure than is proposed here. To overcome this problem, the encoder may elect to define additional elements using the customization methods discussed in *24.3. Customization*. The TEI class system makes such customizations simpler to effect and easier to use in interchange.

These classes are specific to parts of the header:

`model.applicationLike` groups elements used to record application-specific information about a document in its header.

| | |
|---|---|
| application | provides information about an application which has acted upon the document. |

`model.availabilityPart` groups elements such as licences and paragraphs of text which may appear as part of an availability statement.

| | |
|---|---|
| licence | contains information about a licence or other legal agreement applicable to the text. |

`model.catDescPart` groups component elements of the TEI header Category Description.

| | |
|---|---|
| textDesc | (text description) provides a description of a text in terms of its situational parameters. |

`model.editorialDeclPart` groups elements which may be used inside <editorialDecl> and appear multiple times.

| | |
|---|---|
| correction | (correction principles) states how and under what circumstances corrections have been made in the text. |
| hyphenation | (hyphenation) summarizes the way in which hyphenation in a source text has been treated in an encoded version of it. |
| interpretation | (interpretation) describes the scope of any analytic or interpretive information added to the text in addition to the transcription. |
| normalization | (normalization) indicates the extent of normalization or regularization of the original source carried out in converting it to electronic form. |
| punctuation | specifies editorial practice adopted with respect to punctuation marks in the original. |
| quotation | (quotation) specifies editorial practice adopted with respect to quotation marks in the original. |
| segmentation | (segmentation) describes the principles according to which the text has been segmented, for example into sentences, tone-units, graphemic strata, etc. |

| | |
|---|---|
| stdVals | (standard values) specifies the format used when standardized date or number values are supplied. |

`model.encodingDescPart` groups elements which may be used inside <encodingDesc> and appear multiple times.

| | |
|---|---|
| appInfo | (application information) records information about an application which has edited the TEI file. |
| charDecl | (character declarations) provides information about nonstandard characters and glyphs. |
| classDecl | (classification declarations) contains one or more taxonomies defining any classificatory codes used elsewhere in the text. |
| constraintDecl | (constraint declaration) contains declarations pertaining to formal constraints expressed elsewhere in <constraintSpec> elements |
| editorialDecl | (editorial practice declaration) provides details of editorial principles and practices applied during the encoding of a text. |
| fsdDecl | (feature system declaration) provides a feature system declaration comprising one or more feature structure declarations or feature structure declaration links. |
| geoDecl | (geographic coordinates declaration) documents the notation and the datum used for geographic coordinates expressed as content of the <geo> element elsewhere within the document. |
| listPrefixDef | (list of prefix definitions) contains a list of definitions of prefixing schemes used in `teidata.pointer` values, showing how abbreviated URIs using each scheme may be expanded into full URIs. |
| metDecl | (metrical notation declaration) documents the notation employed to represent a metrical pattern when this is specified as the value of a *met*, *real*, or *rhyme* attribute on any structural element of a metrical text (e.g. <lg>, <l>, or <seg>). |
| projectDesc | (project description) describes in detail the aim or purpose for which an electronic file was encoded, together with any other relevant information concerning the process by which it was assembled or collected. |
| refsDecl | (references declaration) specifies how canonical references are constructed for this text. |
| samplingDecl | (sampling declaration) contains a prose description of the rationale and methods used in selecting texts, or parts of a text, for inclusion in the resource. |
| schemaRef | (schema reference) describes or points to a related customization or schema file. |
| schemaSpec | (schema specification) generates a TEI-conformant schema and documentation for it. |
| styleDefDecl | (style definition language declaration) specifies the name of the formal language in which style or renditional information is supplied elsewhere in the document. The specific version of the scheme may also be supplied. |
| tagsDecl | (tagging declaration) provides detailed information about the tagging applied to a document. |
| transcriptionDesc | (transcription description) describes the set of transcription conventions used, particularly for spoken material. |
| unitDecl | (unit declarations) provides information about units of measurement that are not members of the International System of Units. |
| variantEncoding | (variant encoding) declares the method used to encode text-critical variants. |

`model.profileDescPart` groups elements which may be used inside <profileDesc> and appear multiple times.

| | |
|---|---|
| abstract | contains a summary or formal abstract prefixed to an existing source document by the encoder. |
| calendarDesc | (calendar description) contains a description of the calendar system used in any dating expression found in the text. |
| correspDesc | (correspondence description) contains a description of the actions related to one act of correspondence. |
| creation | (creation) contains information about the creation of a text. |
| handNotes | contains one or more <handNote> elements documenting the different hands identified within the source texts. |
| langUsage | (language usage) describes the languages, sublanguages, registers, dialects, etc. represented within a text. |
| listTranspose | supplies a list of transpositions, each of which is indicated at some point in a document typically by means of metamarks. |
| particDesc | (participation description) describes the identifiable speakers, voices, or other participants in any kind of text or other persons named or otherwise referred to in a text, edition, or metadata. |
| settingDesc | (setting description) describes the setting or settings within which a language interaction takes place, or other places otherwise referred to in a text, edition, or metadata. |
| textClass | (text classification) groups information which describes the nature or topic of a text in terms of a standard classification scheme, thesaurus, etc. |
| textDesc | (text description) provides a description of a text in terms of its situational parameters. |

`model.teiHeaderPart` groups high level elements which may appear more than once in a TEI header.

| | |
|---|---|
| encodingDesc | (encoding description) documents the relationship between an electronic text and the source or sources from which it was derived. |
| profileDesc | (text-profile description) provides a detailed description of non-bibliographic aspects of a text, specifically the languages and sublanguages used, the situation in which it was produced, the participants and their setting. |
| xenoData | (non-TEI metadata) provides a container element into which metadata in non-TEI formats may be placed. |

`model.sourceDescPart` groups elements which may be used inside <sourceDesc> and appear multiple times.

| | |
|---|---|
| recordingStmt | (recording statement) describes a set of recordings used as the basis for transcription of a spoken text. |
| scriptStmt | (script statement) contains a citation giving details of the script used for a spoken text. |

`model.textDescPart` groups elements used to categorize a text for example in terms of its situational parameters.

| | |
|---|---|
| channel | (primary channel) describes the medium or channel by which a text is delivered or experienced. For a written text, this might be print, manuscript, email, etc.; for a spoken one, radio, telephone, face-to-face, etc. |
| constitution | (constitution) describes the internal composition of a text or text sample, for example as fragmentary, complete, etc. |
| derivation | (derivation) describes the nature and extent of originality of this text. |

| | |
|---|---|
| domain | (domain of use) describes the most important social context in which the text was realized or for which it is intended, for example private vs. public, education, religion, etc. |
| factuality | (factuality) describes the extent to which the text may be regarded as imaginative or non-imaginative, that is, as describing a fictional or a non-fictional world. |
| interaction | (interaction) describes the extent, cardinality and nature of any interaction among those producing and experiencing the text, for example in the form of response or interjection, commentary, etc. |
| preparedness | (preparedness) describes the extent to which a text may be regarded as prepared or spontaneous. |

## 2.2 The File Description

This section describes the <fileDesc> element, which is the first component of the <teiHeader> element.

The bibliographic description of a machine-readable or digital text resembles in structure that of a book, an article, or any other kind of textual object. The file description element of the TEI header has therefore been closely modelled on existing standards in library cataloguing; it should thus provide enough information to allow users to give standard bibliographic references to the electronic text, and to allow cataloguers to catalogue it. Bibliographic citations occurring elsewhere in the header, and also in the text itself, are derived from the same model (on bibliographic citations in general, see further section *3.12. Bibliographic Citations and References*). See further section *2.8. Note for Library Cataloguers*.

The bibliographic description of an electronic text should be supplied by the mandatory <fileDesc> element:

`<fileDesc>` (file description) contains a full bibliographic description of an electronic file.

The <fileDesc> element contains three mandatory elements and four optional elements, each of which is described in more detail in sections *2.2.1. The Title Statement* to *2.2.6. The Notes Statement* below. These elements are listed below in the order in which they must be given within the <fileDesc> element.

`<titleStmt>` (title statement) groups information about the title of a work and those responsible for its content.

`<editionStmt>` (edition statement) groups information relating to one edition of a text.

`<extent>` (extent) describes the approximate size of a text stored on some carrier medium or of some other object, digital or non-digital, specified in any convenient units.

`<publicationStmt>` (publication statement) groups information concerning the publication or distribution of an electronic or other text.

`<seriesStmt>` (series statement) groups information about the series, if any, to which a publication belongs.

`<notesStmt>` (notes statement) collects together any notes providing information about a text additional to that recorded in other parts of the bibliographic description.

`<sourceDesc>` (source description) describes the source(s) from which an electronic text was derived or generated, typically a bibliographic description in the case of a digitized text, or a phrase such as 'born digital' for a text which has no previous existence.

A complete file description containing all possible sub-elements might look like this:

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>
<!-- title of the resource -->
      </title>
    </titleStmt>
    <editionStmt>
      <p>
<!-- information about the edition of the
resource -->
```

```
      </p>
    </editionStmt>
    <extent>
<!-- description of the size of the resource -->
    </extent>
    <publicationStmt>
      <p>
<!-- information about the distribution
of the resource -->
      </p>
    </publicationStmt>
    <seriesStmt>
      <p>
<!-- information about any series to which
the resource belongs -->
      </p>
    </seriesStmt>
    <notesStmt>
      <note>
<!-- notes on other aspects of the resource -->
      </note>
    </notesStmt>
    <sourceDesc>
      <p>
<!-- information about the source from which
the resource was derived -->
      </p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

Of these elements, only the <titleStmt>, <publicationStmt>, and <sourceDesc> are required; the others may be omitted unless considered useful.

### 2.2.1 The Title Statement

The <titleStmt> element is the first component of the <fileDesc> element, and is mandatory:

`<titleStmt>` (title statement) groups information about the title of a work and those responsible for its content.

It contains the title given to the electronic work, together with one or more optional *statements of responsibility* which identify the encoder, editor, author, compiler, or other parties responsible for it:

`<title>` (title) contains a title for any kind of work.

`<author>` (author) in a bibliographic reference, contains the name(s) of an author, personal or corporate, of a work; for example in the same form as that provided by a recognized bibliographic name authority.

`<editor>` contains a secondary statement of responsibility for a bibliographic item, for example the name of an individual, institution or organization, (or of several such) acting as editor, compiler, translator, etc.

`<sponsor>` (sponsor) specifies the name of a sponsoring organization or institution.

`<funder>` (funding body) specifies the name of an individual, institution, or organization responsible for the funding of a project or text.

`<principal>` (principal researcher) supplies the name of the principal researcher responsible for the creation of an electronic text.

`<respStmt>` (statement of responsibility) supplies a statement of responsibility for the intellectual content of a text, edition, recording, or series, where the specialized elements for authors, editors, etc. do not suffice or do not apply. May also be used to encode information about individuals or organizations which have played a role in the production or distribution of a bibliographic work.

`<resp>` (responsibility) contains a phrase describing the nature of a person's intellectual responsibility, or an organization's role in the production or distribution of a work.

`<name>` (name, proper noun) contains a proper noun or noun phrase.

The <title> element contains the chief name of the electronic work, including any alternative title or subtitles it may have. It may be repeated, if the work has more than one title (perhaps in different languages) and takes whatever form is considered appropriate by its creator. Where the electronic work is derived from an existing source text, it is strongly recommended that the title for the former should be derived from the latter, but clearly distinguishable from it, for example by the addition of a phrase such as ': an electronic transcription' or 'a digital edition'. This will distinguish the electronic work from the source text in citations and in catalogues which contain descriptions of both types of material.

The electronic work will also have an external name (its 'filename' or 'data set name') or reference number on the computer system where it resides at any time. This name is likely to change frequently, as new copies of the file are made on the computer system. Its form is entirely dependent on the particular computer system in use and thus cannot always easily be transferred from one system to another. Moreover, a given work may be composed of many files. For these reasons, these Guidelines strongly recommend that such names should *not* be used as the <title> for any electronic work.

Helpful guidance on the formulation of useful descriptive titles in difficult cases may be found in chapter 25 of ?) or another national cataloguing code.

The elements <author>, <editor>, <sponsor>, <funder>, and <principal>, are specializations of the more general <respStmt> element. These elements are used to provide the *statements of responsibility* which identify the person(s) responsible for the intellectual or artistic content of an item and any corporate bodies from which it emanates.

Any number of such statements may occur within the title statement. At a minimum, identify the author of the text and (where appropriate) the creator of the file. If the bibliographic description is for a corpus, identify the creator of the corpus. Optionally include also names of others involved in the transcription or elaboration of the text, sponsors, and funding agencies. The name of the person responsible for physical data input need not normally be recorded, unless that person is also intellectually responsible for some aspect of the creation of the file.

Where the person whose responsibility is to be documented is not an author, sponsor, funding body, or principal researcher, the <respStmt> element should be used. This has two subcomponents: a <name> element identifying a responsible individual or organization, and a <resp> element indicating the nature of the responsibility. No specific recommendations are made at this time as to appropriate content for the <resp>: it should make clear the nature of the responsibility concerned, as in the examples below.

Names given may be personal names or corporate names. Give all names in the form in which the persons or bodies wish to be publicly cited. This would usually be the fullest form of the name, including first names.[39]

Examples:

```
<titleStmt>
  <title>Capgrave's Life of St. John Norbert: a
     machine-readable transcription</title>
  <respStmt>
    <resp>compiled by</resp>
    <name>P.J. Lucas</name>
  </respStmt>
</titleStmt>
```

```
<titleStmt>
  <title>Two stories by Edgar Allen Poe: electronic version</title>
  <author>Poe, Edgar Allen (1809-1849)</author>
  <respStmt>
    <resp>compiled by</resp>
```

[39] Agencies compiling catalogues of machine-readable files are recommended to use available authority lists, such as the Library of Congress Name Authority List, for all common personal names.

```
    <name>James D. Benson</name>
  </respStmt>
</titleStmt>
```

```
<titleStmt>
  <title>Yogadarśanam (arthāt
     yogasūtrapūṭhaḥ):
     a digital edition.</title>
  <title>The Yogasūtras of Patañjali:
     a digital edition.</title>
  <funder>Wellcome Institute for the History of Medicine</funder>
  <principal>Dominik Wujastyk</principal>
  <respStmt>
    <name>Wieslaw Mical</name>
    <resp>data entry and proof correction</resp>
  </respStmt>
  <respStmt>
    <name>Jan Hajic</name>
    <resp>conversion to TEI-conformant markup</resp>
  </respStmt>
</titleStmt>
```

### 2.2.2 The Edition Statement

The <editionStmt> element is the second component of the <fileDesc> element. It is optional but recommended.

`<editionStmt>` (edition statement) groups information relating to one edition of a text.

It contains either phrases or more specialized elements identifying the edition and those responsible for it:

`<edition>` (edition) describes the particularities of one edition of a text.

`<respStmt>` (statement of responsibility) supplies a statement of responsibility for the intellectual content of a text, edition, recording, or series, where the specialized elements for authors, editors, etc. do not suffice or do not apply. May also be used to encode information about individuals or organizations which have played a role in the production or distribution of a bibliographic work.

`<name>` (name, proper noun) contains a proper noun or noun phrase.

`<resp>` (responsibility) contains a phrase describing the nature of a person's intellectual responsibility, or an organization's role in the production or distribution of a work.

For printed texts, the word *edition* applies to the set of all the identical copies of an item produced from one master copy and issued by a particular publishing agency or a group of such agencies. A change in the identity of the distributing body or bodies does not normally constitute a change of edition, while a change in the master copy does.

For electronic texts, the notion of a 'master copy' is not entirely appropriate, since they are far more easily copied and modified than printed ones; nonetheless the term *edition* may be used for a particular state of a machine-readable text at which substantive changes are made and fixed. Synonymous terms used in these Guidelines are *version*, *level*, and *release*. The words *revision* and *update*, by contrast, are used for minor changes to a file which do not amount to a new edition.

No simple rule can specify how 'substantive' changes have to be before they are regarded as producing a new edition, rather than a simple update. The general principle proposed here is that the production of a new edition entails a significant change in the intellectual content of the file, rather than its encoding or appearance. The addition of analytic coding to a text would thus constitute a new edition, while automatic conversion from one coded representation to another would not. Changes relating to the character code or physical storage details, corrections of misspellings, simple changes in the arrangement of the contents and changes in the output format do not normally constitute a new edition, whereas the addition of new information (e.g. a linguistic analysis expressed in part-of-speech tagging, sound or graphics, referential links to external data sets) almost always does.

Clearly, there will always be borderline cases and the matter is somewhat arbitrary. The simplest rule is: if you think that your file is a new edition, then call it such. An edition statement is optional for the first release of a computer file; it is mandatory for each later release, though this requirement cannot be enforced by the parser.

Note that *all* changes in a file considered significant, whether or not they are regarded as constituting a new edition or simply a new revision, should be independently noted in the revision description section of the file header (see section *2.6. The Revision Description*).

The <edition> element should contain phrases describing the edition or version, including the word *edition*, *version*, or equivalent, together with a number or date, or terms indicating difference from other editions such as *new edition*, *revised edition* etc. Any dates that occur within the edition statement should be marked with the <date> element. The *n* attribute of the <edition> element may be used as elsewhere to supply any formal identification (such as a version number) for the edition.

One or more <respStmt> elements may also be used to supply statements of responsibility for the edition in question. These may refer to individuals or corporate bodies and can indicate functions such as that of a reviser, or can name the person or body responsible for the provision of supplementary matter, of appendices, etc., in a new edition. For further detail on the <respStmt> element, see section *3.12. Bibliographic Citations and References*.

Some examples follow:

```
<editionStmt>
  <edition  n="P2">Second draft, substantially
     extended, revised, and corrected.</edition>
</editionStmt>
```

```
<editionStmt>
  <edition>Student's edition, <date>June 1987</date>
  </edition>
  <respStmt>
    <resp>New annotations by</resp>
    <name>George Brown</name>
  </respStmt>
</editionStmt>
```

### 2.2.3 Type and Extent of File

The <extent> element is the third component of the <fileDesc> element. It is optional.

`<extent>` (extent) describes the approximate size of a text stored on some carrier medium or of some other object, digital or non-digital, specified in any convenient units.

For printed books, information about the carrier, such as the kind of medium used and its size, are of great importance in cataloguing procedures. The print-oriented rules for bibliographic description of an item's medium and extent need some re-interpretation when applied to electronic media. An electronic file exists as a distinct entity quite independently of its carrier and remains the same intellectual object whether it is stored on a magnetic tape, a CD-ROM, a set of floppy disks, or as a file on a mainframe computer. Since, moreover, these Guidelines are specifically aimed at facilitating transparent document storage and interchange, any purely machine-dependent information should be irrelevant as far as the file header is concerned.

This is particularly true of information about *file-type* although library-oriented rules for cataloguing often distinguish two types of computer file: 'data' and 'programs'. This distinction is quite difficult to draw in some cases, for example, hypermedia or texts with built in search and retrieval software.

Although it is equally system-dependent, some measure of the size of the computer file may be of use for cataloguing and other practical purposes. Because the measurement and expression of file size is fraught with difficulties, only very general recommendations are possible; the element <extent> is provided for this purpose. It contains a phrase indicating the size or approximate size of the computer file in one of the following ways:

- in bytes of a specified length (e.g. '4000 16-bit bytes')
- as falling within a range of categories, for example:

  - less than 1 Mb
  - between 1 Mb and 5 Mb
  - between 6 Mb and 10 Mb
  - over 10 Mb
- in terms of any convenient logical units (for example, words or sentences, citations, paragraphs)
- in terms of any convenient physical units (for example, blocks, disks, tapes)

The use of standard abbreviations for units of quantity is recommended where applicable, here as elsewhere (see `http://physics.nist.gov/cuu/Units/binary.html`).

Examples:

```
<extent>between 1 and
2 Mb</extent>
<extent>4.2 MiB</extent>
<extent>4532 bytes</extent>
<extent>3200 sentences</extent>
<extent>Five 90 mm High Density Diskettes</extent>
```

The <measure> element and its attributes may be used to supply machine-tractable or normalized versions of the size or sizes given, as in the following example:

```
<extent>
  <measure  unit="MiB"  quantity="4.2">About four megabytes</measure>
  <measure  unit="pages"  quantity="245">245 pages of source
     material</measure>
</extent>
```

Note that when more than one <measure> is supplied in a single <extent>, the implication is that all the measurements apply to the whole resource.

### 2.2.4 Publication, Distribution, Licensing, etc.

The <publicationStmt> element is the fourth component of the <fileDesc> element and is mandatory. Its function is to name the agency by which a resource is made available (for example, a publisher or distributor) and to supply any additional information about the way in which it is made available such as licensing conditions, identifying numbers, etc.

`<publicationStmt>` (publication statement) groups information concerning the publication or distribution of an electronic or other text.

It may contain either a simple prose description organized as one or more paragraphs, or the more specialised elements described below.

A structured publication statement must begin with one of the following elements:

`<publisher>` (publisher) provides the name of the organization responsible for the publication or distribution of a bibliographic item.

`<distributor>` (distributor) supplies the name of a person or other agency responsible for the distribution of a text.

`<authority>` (release authority) supplies the name of a person or other agency responsible for making a work available, other than a publisher or distributor.

These elements form the `model.publicationStmtPart.agency` class; if the agency making the resource available is unknown, but other structured information about it is available, an explicit statement such as 'publisher unknown' should be used.

The *publisher* is the person or institution by whose authority a given edition of the file is made public. The *distributor* is the person or institution from whom copies of the text may be obtained. Where a text is not considered

formally published, but is nevertheless made available for circulation by some individual or organization, this person or institution is termed the *release authority*.

Whichever of these elements is chosen, it may be followed by one or more of the following elements, which together form the `model.publicationStmtPart.detail` class

`<pubPlace>` (publication place) contains the name of the place where a bibliographic item was published.

`<address>` (address) contains a postal address, for example of a publisher, an organization, or an individual.

`<idno>` (identifier) supplies any form of identifier used to identify some object, such as a bibliographic item, a person, a title, an organization, etc. in a standardized way.

`@type` categorizes the identifier, for example as an ISBN, Social Security number, etc.

`<availability>` (availability) supplies information about the availability of a text, for example any restrictions on its use or distribution, its copyright status, any licence applying to it, etc.

`@status` (status) supplies a code identifying the current availability of the text.

`<date>` (date) contains a date in any format.

`<licence>` contains information about a licence or other legal agreement applicable to the text.

Here is a simple example:

```
<publicationStmt>
  <publisher>Oxford University Press</publisher>
  <pubPlace>Oxford</pubPlace>
  <date>1989</date>
  <idno  type="ISBN">0-19-254705-4</idno>
  <availability>
    <p>Copyright 1989, Oxford University Press</p>
  </availability>
</publicationStmt>
```

The `model.publicationStmtPart.detail` elements all supply additional information relating to the the publisher, distributor, or release authority immediately preceding them. In the following example, Benson is identified as responsible for distribution of some resource at the date and place cited:

```
<publicationStmt>
  <authority>James D. Benson</authority>
  <pubPlace>London</pubPlace>
  <date>1994</date>
</publicationStmt>
```

A resource may have (for example) both a publisher and a distributor, or more than one publisher each using different identifiers for the same resource, and so on. For this reason, the sequence of at least one `model.publicationStmtPart.agency` element followed by zero or more `model.publicationStmtPart.detail` elements may be repeated as often as necessary.

The following example shows a resource published by one agency (Sigma Press) at one address and date, which is also distributed by another (Oxford Text Archive), with a specified identifier and a different date:

```
<publicationStmt>
  <publisher>Sigma Press</publisher>
  <address>
    <addrLine>21 High Street,</addrLine>
    <addrLine>Wilmslow,</addrLine>
    <addrLine>Cheshire M24 3DF</addrLine>
  </address>
  <date>1991</date>
```

```
  <distributor>Oxford Text Archive</distributor>
  <idno  type="OTA">1256</idno>
  <availability>
    <p>Available with prior consent of depositor for
        purposes of academic research and teaching only.</p>
  </availability>
  <date>1994</date>
</publicationStmt>
```

The <date> element used within <publicationStmt> always refers to the date of publication, first distribution, or initial release. If the text was created at some other date, this may be recorded using the <creation> element within the <profileDesc> element. Other useful dates (such as dates of collection of data) may be given using a note in the <notesStmt> element.

The <availability> element may be used, as above, to provide a simple prose statement of any restrictions concerning the distribution of the resource. Alternatively, a more formal statement of the licensing conditions applicable may be provided using the <licence> element:

```
<publicationStmt>
  <publisher>University of Victoria Humanities Computing and Media Centre</publisher>
  <pubPlace>Victoria, BC</pubPlace>
  <date>2011</date>
  <availability  status="restricted">
    <licence  target="http://creativecommons.org/licenses/by-sa/3.0/"> Distributed
under a Creative Commons Attribution-ShareAlike 3.0 Unported License
    </licence>
  </availability>
</publicationStmt>
```

Note here the use of the *target* attribute to point to a location from which the licence document itself may be obtained. Alternatively, the licence document may simply be contained within the <licence> element.

### 2.2.5 The Series Statement

The <seriesStmt> element is the fifth component of the <fileDesc> element and is optional.

`<seriesStmt>` (series statement) groups information about the series, if any, to which a publication belongs.

In bibliographic parlance, a *series* may be defined in one of the following ways:

- A group of separate items related to one another by the fact that each item bears, in addition to its own title proper, a collective title applying to the group as a whole. The individual items may or may not be numbered.
- Each of two or more volumes of essays, lectures, articles, or other items, similar in character and issued in sequence.
- A separately numbered sequence of volumes within a series or serial.

A <seriesStmt> element may contain a prose description or one or more of the following more specific elements:

`<title>` (title) contains a title for any kind of work.

`<idno>` (identifier) supplies any form of identifier used to identify some object, such as a bibliographic item, a person, a title, an organization, etc. in a standardized way.

`<respStmt>` (statement of responsibility) supplies a statement of responsibility for the intellectual content of a text, edition, recording, or series, where the specialized elements for authors, editors, etc. do not suffice or do not apply. May also be used to encode information about individuals or organizations which have played a role in the production or distribution of a bibliographic work.

`<resp>` (responsibility) contains a phrase describing the nature of a person's intellectual responsibility, or an organization's role in the production or distribution of a work.

`<name>` (name, proper noun) contains a proper noun or noun phrase.

The <idno> may be used to supply any identifying number associated with the item, including both standard numbers such as an ISSN and particular issue numbers. (Arabic numerals separated by punctuation are recommended for this purpose: 6.19.33, for example, rather than VI/xix:33). Its *type* attribute is used to categorize the number further, taking the value ISSN for an ISSN for example. Multiple <seriesStmt> elements may be supplied if the TEI document is associated with more than one series.

Examples:

```
<seriesStmt>
  <title  level="s">Machine-Readable Texts for the Study of
     Indian Literature</title>
  <respStmt>
    <resp>ed. by</resp>
    <name>Jan Gonda</name>
  </respStmt>
  <biblScope  unit="volume">1.2</biblScope>
  <idno  type="ISSN">0 345 6789</idno>
</seriesStmt>
```

### 2.2.6 The Notes Statement

The <notesStmt> element is the sixth component of the <fileDesc> element and is optional. If used, it contains one or more <note> elements, each containing a single piece of descriptive information of the kind treated as 'general notes' in traditional bibliographic descriptions.

`<notesStmt>` (notes statement) collects together any notes providing information about a text additional to that recorded in other parts of the bibliographic description.

`<note>` (note) contains a note or annotation.

Some information found in the notes area in conventional bibliography has been assigned specific elements in these Guidelines; in particular the following items should be tagged as indicated, rather than as general notes:

- the nature, scope, artistic form, or purpose of the file; also the genre or other intellectual category to which it may belong: e.g. 'Text types: newspaper editorials and reportage, science fiction, westerns, and detective stories'. These should be formally described within the <profileDesc> element (section *2.4. The Profile Description*).
- an abstract or summary of the content of a document which has been supplied by the encoder because no such abstract forms part of the content of the source. This should be supplied in the <abstract> element within the <profileDesc> element (section *2.4. The Profile Description*).
- summary description providing a factual, non-evaluative account of the subject content of the file: e.g. 'Transcribes interviews on general topics with native speakers of English in 17 cities during the spring and summer of 1963.' These should also be formally described within the <profileDesc> element (section *2.4. The Profile Description*).
- bibliographic details relating to the source or sources of an electronic text: e.g. 'Transcribed from the Norton facsimile of the 1623 Folio'. These should be formally described in the <sourceDesc> element (section *2.2.7. The Source Description*).
- further information relating to publication, distribution, or release of the text, including sources from which the text may be obtained, any restrictions on its use or formal terms on its availability. These should be placed in the appropriate division of the <publicationStmt> element (section *2.2.4. Publication, Distribution, Licensing, etc.*).
- publicly documented numbers associated with the file: e.g. 'ICPSR study number 1803' or 'Oxford Text Archive text number 1243'. These should be placed in an <idno> element within the appropriate division of the <publicationStmt> element. International Standard Serial Numbers (ISSN), International Standard Book Numbers (ISBN), and other internationally agreed upon standard numbers that uniquely identify an item, should be treated in the same way, rather than as specialized bibliographic notes.

Nevertheless, the <notesStmt> element may be used to record potentially significant details about the file and its features, e.g.:

- dates, when they are relevant to the content or condition of the computer file: e.g. 'manual dated 1983', 'Interview wave I: Apr. 1989; wave II: Jan. 1990'
- names of persons or bodies connected with the technical production, administration, or consulting functions of the effort which produced the file, if these are not named in statements of responsibility in the title or edition statements of the file description: e.g. 'Historical commentary provided by Mark Cohen'
- availability of the file in an additional medium or information not already recorded about the availability of documentation: e.g. 'User manual is loose-leaf in eleven paginated sections'
- language of work and abstract, if not encoded in the <langUsage> element, e.g. 'Text in English with summaries in French and German'
- The unique name assigned to a serial by the International Serials Data System (ISDS), if not encoded in an <idno>
- lists of related publications, either describing the source itself, or concerned with the creation or use of the electronic work, e.g. 'Texts used in ?'

Each such item of information may be tagged using the general-purpose <note> element, which is described in section *3.9. Notes, Annotation, and Indexing*. Groups of notes are contained within the <notesStmt> element, as in the following example:

```
<notesStmt>
  <note>Historical commentary provided by Mark Cohen.</note>
  <note>OCR scanning done at University of Toronto.</note>
</notesStmt>
```

There are advantages, however, to encoding such information with more precise elements elsewhere in the TEI header, when such elements are available. For example, the notes above might be encoded as follows:

```
<titleStmt>
  <title>…</title>
  <respStmt>
    <persName>Mark Cohen</persName>
    <resp>historical commentary</resp>
  </respStmt>
  <respStmt>
    <orgName>University of Toronto</orgName>
    <resp>OCR scanning</resp>
  </respStmt>
</titleStmt>
```

### 2.2.7 The Source Description

The <sourceDesc> element is the seventh and final component of the <fileDesc> element. It is a mandatory element and is used to record details of the source or sources from which a computer file is derived. This might be a printed text or manuscript, another computer file, an audio or video recording of some kind, or a combination of these. An electronic file may also have no source, if what is being catalogued is an original text created in electronic form.

`<sourceDesc>` (source description) describes the source(s) from which an electronic text was derived or generated, typically a bibliographic description in the case of a digitized text, or a phrase such as 'born digital' for a text which has no previous existence.

The <sourceDesc> element may contain little more than a simple prose description, or a brief note stating that the document has no source:

```
<sourceDesc>
  <p>Born digital.</p>
</sourceDesc>
```

Alternatively, it may contain elements drawn from the following three classes:

`model.biblLike` groups elements containing a bibliographic description.

| | |
|---|---|
| bibl | (bibliographic citation) contains a loosely-structured bibliographic citation of which the sub-components may or may not be explicitly tagged. |
| biblFull | (fully-structured bibliographic citation) contains a fully-structured bibliographic citation, in which all components of the TEI file description are present. |
| biblStruct | (structured bibliographic citation) contains a structured bibliographic citation, in which only bibliographic sub-elements appear and in a specified order. |
| listBibl | (citation list) contains a list of bibliographic citations of any kind. |
| msDesc | (manuscript description) contains a description of a single identifiable manuscript or other text-bearing object such as an early printed book. |

`model.sourceDescPart` groups elements which may be used inside <sourceDesc> and appear multiple times.

| | |
|---|---|
| recordingStmt | (recording statement) describes a set of recordings used as the basis for transcription of a spoken text. |
| scriptStmt | (script statement) contains a citation giving details of the script used for a spoken text. |

`model.listLike` groups list-like elements.

| | |
|---|---|
| list | (list) contains any sequence of items organized as a list. |
| listApp | (list of apparatus entries) contains a list of apparatus entries. |
| listEvent | (list of events) contains a list of descriptions, each of which provides information about an identifiable event. |
| listNym | (list of canonical names) contains a list of nyms, that is, standardized names for any thing. |
| listObject | (list of objects) contains a list of descriptions, each of which provides information about an identifiable physical object. |
| listOrg | (list of organizations) contains a list of elements, each of which provides information about an identifiable organization. |
| listPerson | (list of persons) contains a list of descriptions, each of which provides information about an identifiable person or a group of people, for example the participants in a language interaction, or the people referred to in a historical source. |
| listPlace | (list of places) contains a list of places, optionally followed by a list of relationships (other than containment) defined amongst them. |
| listRelation | provides information about relationships identified amongst people, places, and organizations, either informally as prose or as formally expressed relation links. |
| listWit | (witness list) lists definitions for all the witnesses referred to by a critical apparatus, optionally grouped hierarchically. |
| table | (table) contains text displayed in tabular form, in rows and columns. |

These classes make available by default a range of ways of providing bibliographic citations which specify the provenance of the text. For written or printed sources, the source may be described in the same way as any other bibliographic citation, using one of the following elements:

`<bibl>` (bibliographic citation) contains a loosely-structured bibliographic citation of which the sub-components may or may not be explicitly tagged.

`<biblStruct>` (structured bibliographic citation) contains a structured bibliographic citation, in which only bibliographic sub-elements appear and in a specified order.

`<listBibl>` (citation list) contains a list of bibliographic citations of any kind.

These elements are described in more detail in section *3.12. Bibliographic Citations and References*. Using them, a source might be described in very simple terms:

```
<sourceDesc>
  <bibl>The first folio of Shakespeare, prepared by
     Charlton Hinman (The Norton Facsimile, 1968)</bibl>
</sourceDesc>
```

or with more elaboration:

```
<sourceDesc>
  <biblStruct  xml:lang="fr">
    <monogr>
      <author>Eugène Sue</author>
      <title>Martin, l'enfant trouvé</title>
      <title  type="sub">Mémoires d'un valet de chambre</title>
      <imprint>
        <pubPlace>Bruxelles et Leipzig</pubPlace>
        <publisher>C. Muquardt</publisher>
        <date  when="1846">1846</date>
      </imprint>
    </monogr>
  </biblStruct>
</sourceDesc>
```

When the header describes a text derived from some pre-existing TEI-conformant or other digital document, it may be simpler to use the following element, which is designed specifically for documents derived from texts which were 'born digital':

`<biblFull>` (fully-structured bibliographic citation) contains a fully-structured bibliographic citation, in which all components of the TEI file description are present.

For further discussion see section *2.2.8. TEI documents derived from other TEI documents*.

When the module for manuscript description is included in a schema, this class also makes available the following element:

`<msDesc>` (manuscript description) contains a description of a single identifiable manuscript or other text-bearing object such as an early printed book.

This element enables the encoder to record very detailed information about one or more manuscript or analogous sources, as further discussed in *11. Manuscript Description*.

The `model.sourceDescPart` class also makes available additional elements when additional modules are included. For example, when the `spoken` module is included, the <sourceDesc> element may also include the following special-purpose elements, intended for cases where an electronic text is derived from a spoken text rather than a written one:

`<scriptStmt>` (script statement) contains a citation giving details of the script used for a spoken text.

`<recordingStmt>` (recording statement) describes a set of recordings used as the basis for transcription of a spoken text.

Full descriptions of these elements and their contents are given in section *8.2. Documenting the Source of Transcribed Speech*.

A single electronic text may be derived from multiple source documents, in whole or in part. The <sourceDesc> may therefore contain a <listBibl> element grouping together <bibl>, <biblStruct>, or <msDesc> elements for each of the sources concerned. It is also possible to repeat the <sourceDesc> element in such a case. The *decls* attribute described in section *16.3. Associating Contextual Information with a Text* may be used to associate parts of the encoded text with the bibliographic element from which it derives in either case.

The source description may also include lists of names, persons, places, etc. when these are considered to form part of the source for an encoded document. When such information is recorded using the specialized elements

discussed in the `namesdates` module (*14. Names, Dates, People, and Places*), the class `model.listLike` makes available the following elements to hold such information:

`<listNym>` (list of canonical names) contains a list of nyms, that is, standardized names for any thing.

`<listOrg>` (list of organizations) contains a list of elements, each of which provides information about an identifiable organization.

`<listPerson>` (list of persons) contains a list of descriptions, each of which provides information about an identifiable person or a group of people, for example the participants in a language interaction, or the people referred to in a historical source.

`<listPlace>` (list of places) contains a list of places, optionally followed by a list of relationships (other than containment) defined amongst them.

### 2.2.8 TEI documents derived from other TEI documents

When a TEI document (call it B) is derived not from a printed source but from another TEI file (call it A), then the five sections of A's file header will need to be incorporated into the new header for B as indicated below:

**fileDesc** A's file description should be copied into the <sourceDesc> section of B's file description, enclosed within a <biblFull> element

**profileDesc** A's <profileDesc> should be copied into B's, in principle unchanged; it may however be expanded by project-specific information relating to B.

**encodingDesc** A's encoding practice may or (more likely) may not be the same as B's. Since the object of the encoding description is to define the relationship between the current file and its source, in principle only changes in encoding practice between A and B need be documented in B. The relationship between A and its source(s) is then only recoverable from the original header of A. In practice it may be more convenient to create a new complete <encodingDesc> for B based on A's.

**xenoData** B is a new computer file, with a different source than A's source (namely, A). Thus it is unlikely that metadata from other schemes about A or its source can be copied wholesale to B, although there may be similarities.

**revisionDesc** B is a new computer file, and should therefore have a new revision description. If, however, it is felt useful to include some information from A's <revisionDesc>, for example dates of major updates or versions, such information must be clearly marked as relating to A rather than to B.

This concludes the discussion of the <fileDesc> element and its contents.

## 2.3 The Encoding Description

The <encodingDesc> element is the second major subdivision of the TEI header. It specifies the methods and editorial principles which governed the transcription or encoding of the text in hand and may also include sets of coded definitions used by other components of the header. Though not formally required, its use is highly recommended.

`<encodingDesc>` (encoding description) documents the relationship between an electronic text and the source or sources from which it was derived.

The encoding description may contain any combination of paragraphs of text, marked up using the <p> element, along with more specialized elements taken from the `model.encodingDescPart` class. By default, this class makes available the following elements:

`<projectDesc>` (project description) describes in detail the aim or purpose for which an electronic file was encoded, together with any other relevant information concerning the process by which it was assembled or collected.

`<samplingDecl>` (sampling declaration) contains a prose description of the rationale and methods used in selecting texts, or parts of a text, for inclusion in the resource.

`<editorialDecl>` (editorial practice declaration) provides details of editorial principles and practices applied during the encoding of a text.

`<tagsDecl>` (tagging declaration) provides detailed information about the tagging applied to a document.

`<styleDefDecl>` (style definition language declaration) specifies the name of the formal language in which style or renditional information is supplied elsewhere in the document. The specific version of the scheme may also be supplied.

`<refsDecl>` (references declaration) specifies how canonical references are constructed for this text.

`<classDecl>` (classification declarations) contains one or more taxonomies defining any classificatory codes used elsewhere in the text.

`<geoDecl>` (geographic coordinates declaration) documents the notation and the datum used for geographic coordinates expressed as content of the <geo> element elsewhere within the document.

`<unitDecl>` (unit declarations) provides information about units of measurement that are not members of the International System of Units.

`<schemaSpec>` (schema specification) generates a TEI-conformant schema and documentation for it.

`<schemaRef>` (schema reference) describes or points to a related customization or schema file.

Each of these elements is further described in the appropriate section below. Other modules have the ability to extend this class; examples are noted in section *2.3.12. Module-Specific Declarations*

### 2.3.1 The Project Description

The <projectDesc> element may be used to describe, in prose, the purpose for which a digital resource was created, together with any other relevant information concerning the process by which it was assembled or collected. This is of particular importance for corpora or miscellaneous collections, but may be of use for any text, for example to explain why one kind of encoding practice has been followed rather than another.

`<projectDesc>` (project description) describes in detail the aim or purpose for which an electronic file was encoded, together with any other relevant information concerning the process by which it was assembled or collected.

For example:

```
<encodingDesc>
  <projectDesc>
    <p>Texts collected for use in the
        Claremont Shakespeare Clinic, June 1990.</p>
  </projectDesc>
</encodingDesc>
```

### 2.3.2 The Sampling Declaration

The <samplingDecl> element may be used to describe, in prose, the rationale and methods used in selecting texts, or parts of text, for inclusion in the resource.

`<samplingDecl>` (sampling declaration) contains a prose description of the rationale and methods used in selecting texts, or parts of a text, for inclusion in the resource.

It should include information about such matters as

- the size of individual samples
- the method or methods by which they were selected
- the underlying population being sampled
- the object of the sampling procedure used

but is not restricted to these.

```
<samplingDecl>
  <p>Samples of 2000 words taken from the beginning of the text.</p>
</samplingDecl>
```

It may also include a simple description of any parts of the source text included or excluded.

```
<samplingDecl>
  <p>Text of stories only has been transcribed. Pull quotes, captions,
     and advertisements have been silently omitted. Any mathematical
     expressions requiring symbols not present in the ISOnum or ISOpub
     entity sets have been omitted, and their place marked with a GAP
     element.</p>
</samplingDecl>
```

A sampling declaration which applies to more than one text or division of a text need not be repeated in the header of each such text. Instead, the *decls* attribute of each text (or subdivision of the text) to which the sampling declaration applies may be used to supply a cross-reference to it, as further described in section *16.3. Associating Contextual Information with a Text*.

### 2.3.3 The Editorial Practices Declaration

The <editorialDecl> element is used to provide details of the editorial practices applied during the encoding of a text.

`<editorialDecl>` (editorial practice declaration) provides details of editorial principles and practices applied during the encoding of a text.

It may contain a prose description only, or one or more of a set of specialized elements, members of the TEI `model.editorialDeclPart` class. Where an encoder wishes to record an editorial policy not specified above, this may be done by adding a new element to this class, using the mechanisms discussed in chapter *24.3. Customization*.

Some of these policy elements carry attributes to support automated processing of certain well-defined editorial decisions; all of them contain a prose description of the editorial principles adopted with respect to the particular feature concerned. Examples of the kinds of questions which these descriptions are intended to answer are given in the list below.

`<correction>` (correction principles) states how and under what circumstances corrections have been made in the text.

`@status` indicates the degree of correction applied to the text.

`@method` indicates the method adopted to indicate corrections within the text.

Was the text corrected during or after data capture? If so, were corrections made silently or are they marked using the tags described in section *3.5. Simple Editorial Changes*? What principles have been adopted with respect to omissions, truncations, dubious corrections, alternate readings, false starts, repetitions, etc.?

`<normalization>` (normalization) indicates the extent of normalization or regularization of the original source carried out in converting it to electronic form.

`@source [att.global.source]` specifies the source from which some aspect of this element is drawn.

`@method` indicates the method adopted to indicate normalizations within the text.

Was the text normalized, for example by regularizing any non-standard spellings, dialect forms, etc.? If so, were normalizations performed silently or are they marked using the tags described in section *3.5. Simple Editorial Changes*? What authority was used for the regularization? Also, what principles were used when normalizing numbers to provide the standard values for the *value* attribute described in section *3.6.3. Numbers and Measures* and what format used for them?

`<punctuation>` specifies editorial practice adopted with respect to punctuation marks in the original.

`@marks` indicates whether or not punctation marks have been retained as content within the text.

`@placement` indicates the positioning of punctuation marks that are associated with marked up text as being encoded within the element surrounding the text or immediately before or after it.

Are punctuation marks present in the original source retained? Are they identified with the element <pc>, or implied by markup? If retained, how are they placed with respect to related elements? For example, do commas and periods appear inside or outside elements marking phrases and sentences?

`<quotation>` (quotation) specifies editorial practice adopted with respect to quotation marks in the original.

`@marks` (quotation marks) indicates whether or not quotation marks have been retained as content within the text.

How were quotation marks processed? Are apostrophes and quotation marks distinguished? How? Are quotation marks retained as content in the text or replaced by markup? Are there any special conventions regarding for example the use of single or double quotation marks when nested? Is the file consistent in its practice or has this not been checked? See section *3.3.3. Quotation* for discussion of ways in which quotation marks may be encoded.

`<hyphenation>` (hyphenation) summarizes the way in which hyphenation in a source text has been treated in an encoded version of it.

`@eol` (end-of-line) indicates whether or not end-of-line hyphenation has been retained in a text.

Does the encoding distinguish 'soft' and 'hard' hyphens? What principle has been adopted with respect to end-of-line hyphenation where source lineation has not been retained? Have soft hyphens been silently removed, and if so what is the effect on lineation and pagination? See section *3.2.2. Hyphenation* for discussion of ways in which hyphenation may be encoded.

`<segmentation>` (segmentation) describes the principles according to which the text has been segmented, for example into sentences, tone-units, graphemic strata, etc.

How is the text segmented? If <s> or <seg> segmentation units have been used to divide up the text for analysis, how are they marked and how was the segmentation arrived at?

`<stdVals>` (standard values) specifies the format used when standardized date or number values are supplied.

In most cases, attributes bearing standardized values (such as the *when* or *when-iso* attribute on dates) should conform to a defined W3C or ISO datatype. In cases where this is not appropriate, this element may be used to describe the standardization methods underlying the values supplied.

`<interpretation>` (interpretation) describes the scope of any analytic or interpretive information added to the text in addition to the transcription.

Has any analytic or 'interpretive' information been provided—that is, information which is felt to be non-obvious, or potentially contentious? If so, how was it generated? How was it encoded? If feature-structure analysis has been used, are <fsdDecl> elements (section *19.11. Feature System Declaration*) present?

Any information about the editorial principles applied not falling under one of the above headings should be recorded in a distinct list of items. Experience shows that a full record should be kept of decisions relating to editorial principles and encoding practice, both for future users of the text and for the project which produced the text in the first instance. Some simple examples follow:

```
<editorialDecl>
  <segmentation>
    <p>
      <gi>s</gi> elements mark orthographic sentences and
         are numbered sequentially
         within their parent <gi>div</gi> element
    </p>
  </segmentation>
  <interpretation>
    <p>The part of speech analysis applied throughout section 4 was
         added by hand and has not been checked.</p>
  </interpretation>
  <correction>
    <p>Errors in transcription controlled by using the
         WordPerfect spelling checker.</p>
  </correction>
  <normalization  source="http://szotar.sztaki.hu/webster/">
    <p>All words converted to Modern American spelling following
         Websters 9th Collegiate dictionary.</p>
  </normalization>
  <quotation  marks="all">
    <p>All opening quotation marks represented by entity reference
    <ident  type="ge">odq</ident>; all closing quotation marks
         represented by entity reference <ident  type="ge">cdq</ident>.</p>
  </quotation>
</editorialDecl>
```

An editorial practices declaration which applies to more than one text or division of a text need not be repeated in the header of each such text. Instead, the *decls* attribute of each text (or subdivision of the text) to which it applies may be used to supply a cross-reference to it, as further described in section *16.3. Associating Contextual Information with a Text*.

### 2.3.4 The Tagging Declaration

The <tagsDecl> element is used to record the following information about the tagging used within a particular document:

- the namespace to which elements appearing within the transcribed text belong.
- how often particular elements appear within the text, so that a recipient can validate the integrity of a text during interchange.
- any comment relating to the usage of particular elements not specified elsewhere in the header.
- a default rendition applicable to all instances of an element.

This information is conveyed by the following elements:

`<rendition>` (rendition) supplies information about the rendition or appearance of one or more elements in the source text.

`@selector` contains a selector or series of selectors specifying the elements to which the contained style description applies, expressed in the language specified in the *scheme* attribute.

`att.styleDef` provides attributes to specify the name of a formal definition language used to provide formatting or rendition information.

`@scheme` identifies the language used to describe the rendition.

`@schemeVersion` supplies a version number for the style language provided in *scheme*.

`<namespace>` (namespace) supplies the formal name of the namespace to which the elements documented by its children belong.

`<tagUsage>` (element usage) documents the usage of a specific element within a specified document.

The <tagsDecl> element is descriptive, rather than prescriptive: if used, it simply documents practice in the TEI document containing it. The elements constituting a TEI customization file (discussed in chapter *23. Documentation Elements*) by contrast document expected practice in a number of documents, and may thus be used prescriptively. If there is an inconsistency between the actual state of a document and what is documented by its <tagsDecl>, then the latter should be corrected. If there is an inconsistency between a document and what is required by the customization file, or a schema derived from it, then it will usually be the document that requires correction.

The <tagsDecl> element consists of an optional sequence of <rendition> elements, each of which must bear a unique identifier, followed by an optional sequence of one or more <namespace> elements, each of which contains a series of <tagUsage> elements, up to one for each element type from that namespace occurring within the associated <text> element. Note that these <tagUsage> elements must be nested within a <namespace> element, and cannot appear directly within the <tagsDecl> element.

#### 2.3.4.1 Rendition

The <rendition> element allows the encoder to specify how one or more elements are rendered in the original source in any of the following ways:

- using an informal prose description
- using a standard stylesheet language such as CSS or XSL-FO
- using a project-defined formal language

One or more such specifications may be associated with elements of a document in two ways:

- the *selector* attribute on any <rendition> element may be used to select a collection of elements to which it applies
- the global *rendition* attribute may be used on any element to indicate its rendition, overriding or complementing any supplied default value

The global *rend* and *style* attributes may also be used to describe the rendering of an element. See further *1.3.1.1.3. Rendition Indicators*.

The content of a <rendition> element may describe the appearance of the source material using prose, a project-defined formal language, or any standard languages such as the Cascading Stylesheet Language (?) or the XML vocabulary for specifying formatting semantics which forms a part of the W3C's Extensible Stylesheet Language (?). A <styleDefDecl> element (*2.3.5. The Default Style Definition Language Declaration*) may be supplied within the <encodingDesc> to specify which of these applies by default, and it may be overridden for one or more specific <rendition> elements using the *scheme* attribute.

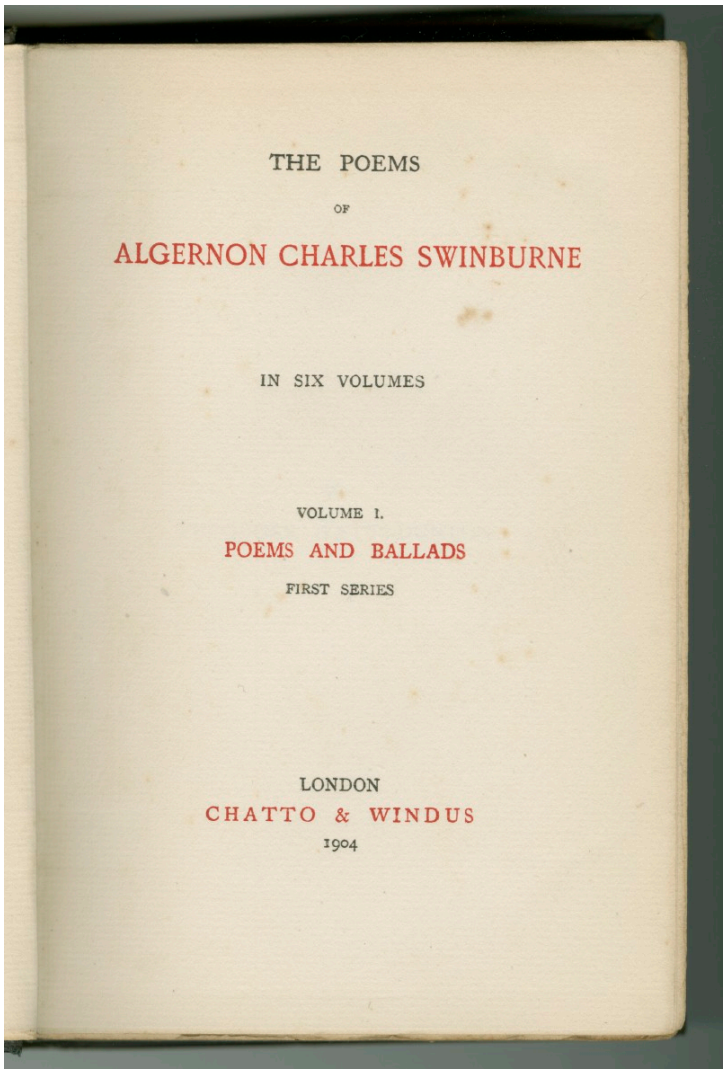
The recommended way to indicate a default rendition on one or more elements is to use the *selector* attribute together with the *scheme* attribute on <rendition>. For example, suppose that all paragraphs in the <front> of a text appear in a large font, with significant top and bottom margins, while paragraphs in the main <body> are in regular font size and have no top and bottom margins. The use of *selector* together with *scheme* provides an efficient way to specify the distinct styling for distinct contexts of the paragraph by means of CSS selectors:

```
<rendition scheme="css" selector="front p">
font-size: 110%;
margin-top: 0.5em;
margin-bottom: 0.5em;
</rendition>
```

```
<rendition scheme="css" selector="body p">
font-size: 100%;
margin-top: 0;
```

```
margin-bottom: 0;
</rendition>
```

In the following extended example we consider how to capture the appearance of a typical early 20th century titlepage, such as that in the following figure: Elements for the encoding of the information on a titlepage are




presented in *4.6. Title Pages*; here we consider how we might go about encoding some of the visual information as well, using the <rendition> element and its corresponding attributes.

First we define a rendition element for each aspect of the source page rendition that we wish to retain. Details of CSS are given in **?**; we use it here simply to provide a vocabulary with which to describe such aspects as font size and style, letter and line spacing, colour, etc. Note that the purpose of this encoding is to describe the original, rather than specify how it should be reproduced, although the two are obviously closely linked.

```
<styleDefDecl  scheme="css"
  schemeVersion="2.1"/>
<!-- ... -->
<tagsDecl>
  <rendition  xml:id="center">text-align: center;</rendition>
```

```
  <rendition  xml:id="small">font-size: small;</rendition>
  <rendition  xml:id="large">font-size: large;</rendition>
  <rendition  xml:id="x-large">font-size: x-large;</rendition>
  <rendition  xml:id="xx-large">font-size: xx-large</rendition>
  <rendition  xml:id="expanded">letter-spacing: +3pt;</rendition>
  <rendition  xml:id="x-space">line-height: 150%;</rendition>
  <rendition  xml:id="xx-space">line-height: 200%;</rendition>
  <rendition  xml:id="red">color: red;</rendition>
</tagsDecl>
```

The global *rendition* attribute can now be used to specify on any element which of the above rendition features apply to it. For example, a title page might be encoded as follows:

```
<titlePage>
  <docTitle  rendition="#center #x-space">
    <titlePart>
      <lb/>
      <hi  rendition="#x-large">THE POEMS</hi>
      <lb/>
      <hi  rendition="#small">OF</hi>
      <lb/>
      <hi  rendition="#red #xx-large">ALGERNON CHARLES SWINBURNE</hi>
      <lb/>
      <hi  rendition="#large #xx-space">IN SIX VOLUMES</hi>
    </titlePart>
    <titlePart  rendition="#xx-space">
      <lb/> VOLUME I.
    <lb/>
      <hi  rendition="#red #x-large">POEMS AND BALLADS</hi>
      <lb/>
      <hi  rendition="#x-space">FIRST SERIES</hi>
    </titlePart>
  </docTitle>
  <docImprint  rendition="#center">
    <lb/>
    <pubPlace  rendition="#xx-space">LONDON</pubPlace>
    <lb/>
    <publisher  rendition="#red #expanded">CHATTO & WINDUS</publisher>
    <lb/>
    <docDate  when="1904"  rendition="#small">1904</docDate>
  </docImprint>
</titlePage>
```

When CSS is used as the style definition language, the *scope* attribute may be used to specify CSS pseudo-elements. These pseudo-elements are used to specify styling applicable to only a portion of the given text. For example, the `first-letter` pseudo-element defines styling to be applied to the first letter in the targeted element, while the `before` and `after` pseudo-elements can be used often in conjunction with the ”content” property to add additional characters which need to be added before or after the element content to make it more closely resemble the appearance of the source.

For example, assuming that a text has been encoded using the <q> element to enclose passages in quotation marks, but the quotation marks themselves have been routinely omitted from the encoding, a set of renditions such as the following:

```
<rendition  xml:id="quoteBefore"
  scheme="css"  scope="before">content:
'“';</rendition>
<rendition  xml:id="quoteAfter"  scheme="css"
  scope="after">content:
'”';</rendition>
```

might be used to predefine pseudo-elements `quoteBefore` and `quoteAfter`. Where a <q> element is actually rendered in the source with initial and final quotation marks, it may then be encoded as follows:

```
<q  rendition="#quoteBefore #quoteAfter">Four score and seven years
ago...</q>
```

#### 2.3.4.2 Tag Usage

As noted above, each <namespace> element, if present, should contain up to one occurrence of a <tagUsage> element for each element type from the given namespace that occurs within the outermost <text> element associated with the <teiHeader> in which it appears.[40] The <tagUsage> element may be used to supply a count of the number of occurrences of this element within the text, which is given as the value of its *occurs* attribute. It may also be used to hold any additional usage information, which is supplied as running prose within the element itself.

For example:

```
<tagUsage  gi="hi"  occurs="28"> Used only to mark English words italicised in the copy
text.
</tagUsage>
```

This indicates that the <hi> element appears a total of 28 times in the <text> element in question, and that the encoder has used it to mark italicised English words only.

The *withId* attribute may optionally be used to specify how many of the occurrences of the element in question bear a value for the global *xml:id* attribute, as in the following example:

```
<tagUsage  gi="pb"  occurs="321"  withId="321"> Marks page beginnings in the York
(1734) edition only </tagUsage>
```

This indicates that the <pb> element occurs 321 times, on each of which an identifier is provided.

The content of the <tagUsage> element is not susceptible of automatic processing. It should not therefore be used to hold information for which provision is already made by other components of the encoding description. A TEI-conformant document is not required to provide any <tagUsage> elements or *occurs* attributes, but if it does, then the counts provided must correspond with the number of such elements present in the associated <text>.

### 2.3.5 The Default Style Definition Language Declaration

The content of the <rendition> element, the value of its *selector* attribute, and the value of the *style* attribute are expressed using one of a small number of formally defined style definition languages. For ease of processing, it is strongly recommended to use a single such language throughout an encoding project, although the TEI system permits a mixture.

The element <styleDefDecl>, a sibling of the <tagsDecl> element, is used to supply the name of the default style definition language. The name is supplied as the value of the *scheme* attribute and may take any of the following values:

[40] In the case of a TEI corpus (*16. Language Corpora*), a <tagsDecl> in a corpus header will describe tag usage across the whole corpus, while one in an individual text header will describe tag usage for the individual text concerned.

**free** Informal free text description

**css** Cascading Stylesheet Language

**xslfo** Extensible Stylesheet Language Formatting Objects

**other** A user-defined formal description language

. The *schemeVersion* attribute may be used to supply the precise version of the style definition language used, and the content of this element, if any, may supply additional information.

When the *style* attribute is used, its value must always be expressed using whichever default style definition language is in force. If more than one occurrence of the <styleDefDecl> is provided, there will be more than one default available, and the *decls* attribute must be used to select which is applicable in a given context, as discussed in section *16.3. Associating Contextual Information with a Text*.

### 2.3.6 The Reference System Declaration

The <refsDecl> element is used to document the way in which any standard referencing scheme built into the encoding works.

`<refsDecl>` (references declaration) specifies how canonical references are constructed for this text.

It may contain either a series of prose paragraphs or the following specialized elements:

`<citeStructure>` (citation structure) declares a structure and method for citing the current document.

`<cRefPattern>` (canonical reference pattern) specifies an expression and replacement pattern for transforming a canonical reference into a URI.

`<refState>` (reference state) specifies one component of a canonical reference defined by the milestone method.

`att.patternReplacement` provides attributes for regular-expression matching and replacement.

- `@matchPattern` specifies a regular expression against which the values of other attributes can be matched.
- `@replacementPattern` specifies a 'replacement pattern', that is, the skeleton of a relative or absolute URI containing references to groups in the *matchPattern* which, once subpattern substitution has been performed, complete the URI.

`att.citeStructurePart` provides attributes for selecting particular elements within a document.

- `@use` (use) supplies an XPath selection pattern using the syntax defined in **?**. The XPath pattern is relative to the context given in *match*, which will either be a sibling attribute in the case of <citeStructure> or on the parent <citeStructure> in the case of <citeData>.

Note that not all possible referencing schemes are equally easily supported by current software systems. A choice must be made between the convenience of the encoder and the likely efficiency of the particular software applications envisaged, in this context as in many others. For a more detailed discussion of referencing systems supported by these Guidelines, see section *3.11. Reference Systems* below.

A referencing scheme may be described in one of four ways using this element:

- as a prose description
- as nested set of citation structure declarations
- as a series of pairs of regular expressions and XPaths
- as a concatenation of sequentially organized *milestone*s

Each method is described in more detail below. Only one method can be used within a single <refsDecl> element.

More than one <refsDecl> element can be included in the header if more than one canonical reference scheme is to be used in the same document, but the current proposals do not check for mutual inconsistency.

#### 2.3.6.1 Prose Method

The referencing scheme may be specified within the <refsDecl> by a simple prose description. Such a description should indicate which elements carry identifying information, and whether this information is represented as attribute values or as content. Any special rules about how the information is to be interpreted when reading or generating a reference string should also be specified here. Such a prose description cannot be processed automatically, and this method of specifying the structure of a canonical reference system is therefore not recommended for automatic processing.

For example:

```
<refsDecl>
  <p>The <att>n</att> attribute of each text in this corpus carries a
     unique identifying code for the whole text. The title of the text is
     held as the content of the first <gi>head</gi> element within each
     text. The <att>n</att> attribute on each <gi>div1</gi> and
  <gi>div2</gi> contains the canonical reference for each such
     division, in the form 'XX.yyy', where XX is the book number in Roman
     numerals, and yyy the section number in arabic. Line beginnings are
     marked by empty <gi>lb</gi> elements, each of which includes the
     through line number in Casaubon's edition as the value of its
  <gi>n</gi> attribute.</p>
  <p>The through line number and the text identifier uniquely identify
     any line. A canonical reference may be made up by concatenating the
  <gi>n</gi> values from the <gi>text</gi>, <gi>div1</gi>, or
  <gi>div2</gi> and calculating the line number within each part.</p>
</refsDecl>
```

#### 2.3.6.2 Search-and-Replace Method

This method often requires a significant investment of effort initially, but permits extremely flexible addressing. For details, see section *17.2.5. Canonical References*.

`<cRefPattern>` (canonical reference pattern) specifies an expression and replacement pattern for transforming a canonical reference into a URI.

#### 2.3.6.3 Milestone Method

This method is appropriate when only 'milestone' tags (see section *3.11.3. Milestone Elements*) are available to provide the required referencing information. It does not provide any abilities which cannot be mimicked by the search-and-replace referencing method discussed in the previous section, but in the cases where it applies, it provides a somewhat simpler notation.

A reference based on milestone tags concatenates the values specified by one or more such tags. Since each tag marks the point at which a value changes, it may be regarded as specifying the *refState* of a variable. A reference declaration using this method therefore specifies the individual components of the canonical reference as a sequence of <refState> elements:

`<refState>` (reference state) specifies one component of a canonical reference defined by the milestone method.

`@delim` (delimiter) supplies a delimiting string following the reference component.

`@length` specifies the fixed length of the reference component.

`att.milestoneUnit` provides attributes to indicate the type of section which is changing at a specific milestone.

`@unit` provides a conventional name for the kind of section changing at this milestone.

For example, the reference 'Matthew 12:34' might be thought of as representing the state of three variables: the `book` variable is in state 'Matthew'; the `chapter` variable is in state '12', and the `verse` variable is in state '34'. If milestone tagging has been used, there should be a tag marking the point in the text at which each of the above 'variables' changes its state.[41] To find 'Matthew 12:34' therefore an application must scan left to right through the text, monitoring changes in the state of each of these three variables as it does so. When all three are simultaneously in the required state, the desired point will have been reached. There may of course be several such points.

The *delim* and *length* attributes are used to specify components of a canonical reference using this method in exactly the same way as for the stepwise method described in the preceding section. The other attributes are used to determine which instances of <milestone> tags in the text are to be checked for state-changes. A state-change is signalled whenever a new <milestone> tag is found with *unit* and, optionally, *ed* attributes identical to those of

[41] On the <milestone> tag itself, what are here referred to as 'variables' are identified by the combination of the *ed* and *unit* attributes.

the <refState> element in question. The value for the new state may be given explicitly by the *n* attribute on the <milestone> element, or it may be implied, if the *n* attribute is not specified.

For example, for canonical references in the form *xx.yyy* where the *xx* represents the page number in the first edition, and *yyy* the line number within this page, a reference system declaration such as the following would be appropriate:

```
<refsDecl>
  <refState ed="first" unit="page" length="2"
    delim="."/>
  <refState ed="first" unit="line" length="3"/>
</refsDecl>
```

This implies that milestone tags of the form

```
<milestone n="II" ed="first" unit="page"/>
<milestone ed="first" unit="line"/>
```

will be found throughout the text, marking the positions at which page and line numbers change. Note that no value has been specified for the *n* attribute on the second milestone tag above; this implies that its value at each state change is monotonically increased. For more detail on the use of milestone tags, see section *3.11.3. Milestone Elements*.

The milestone referencing scheme, though conceptually simple, is not supported by a generic XML parser. Its use places a correspondingly greater burden of verification and accuracy on the encoder.

A reference system declaration which applies to more than one text or division of a text need not be repeated in the header of each such text. Instead, the *decls* attribute of each text (or subdivision of the text) to which the declaration applies may be used to supply a cross-reference to it, as further described in section *16.3. Associating Contextual Information with a Text*.

### 2.3.7 The Classification Declaration

The <classDecl> element is used to group together definitions or sources for any descriptive classification schemes used by other parts of the header or elsewhere in the document. Each such scheme is represented by a <taxonomy> element, which may contain either a simple bibliographic citation, or a definition of the descriptive typology concerned; the following elements are used in defining a descriptive classification scheme:

`<classDecl>` (classification declarations) contains one or more taxonomies defining any classificatory codes used elsewhere in the text.

`<taxonomy>` (taxonomy) defines a typology either implicitly, by means of a bibliographic citation, or explicitly by a structured taxonomy.

`<category>` (category) contains an individual descriptive category, possibly nested within a superordinate category, within a user-defined taxonomy.

`<catDesc>` (category description) describes some category within a taxonomy or text typology, either in the form of a brief prose description or in terms of the situational parameters used by the TEI formal <textDesc>.

The <taxonomy> element has two slightly different, but related, functions. For well-recognized and documented public classification schemes, such as Dewey or other published descriptive thesauri, it contains simply a bibliographic citation indicating where a full description of a particular taxonomy may be found.

```
<taxonomy xml:id="DDC12">
  <bibl>
    <title>Dewey Decimal Classification</title>
    <edition>Abridged Edition 12</edition>
  </bibl>
</taxonomy>
```

For less easily accessible schemes, the <taxonomy> element contains a description of the taxonomy itself as well as an optional bibliographic citation. The description consists of a number of <category> elements, each defining a single category within the given typology. The category is defined by the contents of a nested <catDesc> element, which may contain either a phrase describing the category, or any number of elements from the `model.catDescPart` class. When the corpus module is included in a schema, this class provides the <textDesc> element whose components allow the definition of a text type in terms of a set of 'situational parameters' (see further section *16.2.1. The Text Description*; if the corpus module is not included in a schema, this class is empty and the <catDesc> element may contain only plain text.

If the category is subdivided, each subdivision is represented by a nested <category> element, having the same structure. Categories may be nested to an arbitrary depth in order to reflect the hierarchical structure of the taxonomy. Each <category> element bears a unique *xml:id* attribute, which is used as the target for <catRef> elements referring to it.

```
<taxonomy xml:id="b">
  <bibl>Brown Corpus</bibl>
  <category xml:id="b.a">
    <catDesc>Press Reportage</catDesc>
    <category xml:id="b.a1">
      <catDesc>Daily</catDesc>
    </category>
    <category xml:id="b.a2">
      <catDesc>Sunday</catDesc>
    </category>
    <category xml:id="b.a3">
      <catDesc>National</catDesc>
    </category>
    <category xml:id="b.a4">
      <catDesc>Provincial</catDesc>
    </category>
    <category xml:id="b.a5">
      <catDesc>Political</catDesc>
    </category>
    <category xml:id="b.a6">
      <catDesc>Sports</catDesc>
    </category>
  </category>
  <category xml:id="b.d">
    <catDesc>Religion</catDesc>
    <category xml:id="b.d1">
      <catDesc>Books</catDesc>
    </category>
    <category xml:id="b.d2">
      <catDesc>Periodicals and tracts</catDesc>
    </category>
  </category>
</taxonomy>
```

Linkage between a particular text and a category within such a taxonomy is made by means of the <catRef> element within the <textClass> element, as described in section *2.4.3. The Text Classification*. Where finer-grained analysis is desired, the *ana* attribute on an element in the text could point to a category, as in the following abbreviated example:

```
<taxonomy>
  <category xml:id="poe">
    <catDesc>Poetry</catDesc>
```

```
    <category  xml:id="sonn">
      <catDesc>Sonnet</catDesc>
      <category  xml:id="shakesSonn">
        <catDesc>Shakespearean Sonnet</catDesc>
      </category>
      <category  xml:id="petraSonn">
        <catDesc>Petrarchan Sonnet</catDesc>
      </category>
    </category>
    <category  xml:id="met">
      <catDesc>Metrical Categories</catDesc>
      <category  xml:id="ft">
        <catDesc>Metrical Feet</catDesc>
        <category  xml:id="iamb">
          <catDesc>Iambic</catDesc>
        </category>
        <category  xml:id="troch">
          <catDesc>trochaic</catDesc>
        </category>
      </category>
      <category  xml:id="ftNm">
        <catDesc>Number of feet</catDesc>
        <category  xml:id="penta">
          <catDesc>>Pentameter</catDesc>
        </category>
        <category  xml:id="tetra">
          <catDesc>>Tetrameter</catDesc>
        </category>
      </category>
    </category>
  </category>
</taxonomy>
<!-- elsewhere in document -->
<lg  ana="#shakesSonnet #iamb #penta">
  <l>Shall I compare thee to a summer's day</l>
<!-- ... -->
</lg>
```

Where the taxonomy permits of classification along more than one dimension, more than one category will be referenced by a particular <catRef>, as in the following example, which identifies a text with the sub-categories 'Daily', 'National', and 'Political' within the category 'Press Reportage' as defined above.

```
<catRef  target="#b.a1 #b.a3 #b.a5"/>
```

A single <category> may contain more than one <catDesc> child, when for example the category is described in more than one language, as in the following example:

```
<category  xml:id="lit">
  <catDesc  xml:lang="pl">literatura piękna</catDesc>
  <catDesc  xml:lang="en">fiction</catDesc>
  <category  xml:id="litProza">
    <catDesc  xml:lang="pl">proza</catDesc>
    <catDesc  xml:lang="en">prose</catDesc>
  </category>
```

```
  <category xml:id="litPoezja">
    <catDesc xml:lang="pl">poezja</catDesc>
    <catDesc xml:lang="en">poetry</catDesc>
  </category>
  <category xml:id="litDramat">
    <catDesc xml:lang="pl">dramat</catDesc>
    <catDesc xml:lang="en">drama</catDesc>
  </category>
</category>
```

### 2.3.8 The Geographic Coordinates Declaration

The following element is provided to indicate (within the header of a document, or in an external location) that a particular coordinate notation, or a particular datum, has been employed in a text. The default notation is a string containing two real numbers separated by whitespace, of which the first indicates latitude and the second longitude according to the 1984 World Geodetic System (WGS84).

`<geoDecl>` (geographic coordinates declaration) documents the notation and the datum used for geographic coordinates expressed as content of the <geo> element elsewhere within the document.

`@datum` supplies a commonly used code name for the datum employed.

### 2.3.9 The Unit Declaration

When documents feature units of measurement that are not listed in the International System of Units, the <unitDecl> element may be used in the encoding description to provide definitions and information about their origins and equivalents.

`<unitDecl>` (unit declarations) provides information about units of measurement that are not members of the International System of Units.

The <unitDecl> contains one or more <unitDef> child elements that serve to describe units of measure which may be marked in <unit> elements within the <text>.

`<unitDef>` (unit definition) contains descriptive information related to a specific unit of measurement.

`<unit>` contains a symbol, a word or a phrase referring to a unit of measurement in any kind of formal or informal system.

`<conversion>` defines how to calculate one unit of measure in terms of another.

- `@fromUnit` indicates a source unit of measure that is to be converted into another unit indicated in *toUnit*.
- `@toUnit` the target unit of measurement for a conversion from a source unit referenced in *fromUnit*.
- `@from [att.datable.w3c]` indicates the starting point of the period in standard form, e.g. yyyy-mm-dd.
- `@to [att.datable.w3c]` indicates the ending point of the period in standard form, e.g. yyyy-mm-dd.
- `@notBefore [att.datable.w3c]` specifies the earliest possible date for the event in standard form, e.g. yyyy-mm-dd.
- `@notAfter [att.datable.w3c]` specifies the latest possible date for the event in standard form, e.g. yyyy-mm-dd.
- `@when [att.datable.w3c]` supplies the value of the date or time in a standard form, e.g. yyyy-mm-dd.

`att.formula` provides attributes for defining a mathematical formula.

- `@formula` A *formula* is provided to describe a mathematical calculation such as a conversion between measurement systems.

`att.locatable` provides attributes for referencing locations by pointing to entries in a canonical list of places.

@*where* indicates one or more locations by pointing to a <place> element or other canonical description.

Within the <unitDef>, a <conversion> element may be used to store information relating to conversion between units. The <conversion> element holds a special pair of attributes, *fromUnit* and *toUnit*, which serve to indicate the direction of a calculation from one unit of measure (stored in *fromUnit*) to another (stored in *toUnit*). A mathematical calculation to define the relation between these units may be stored in *formula*, as shown in the following examples. The *formula* attribute takes a value expressed as an XPath expression, for which see **?**. Among other things, this means that division must be expressed with 'div' as opposed to the more common forward slash ('/', U+002F), so as not to be confused with path navigation. The <conversion> element may also use attributes to store dates and date ranges as well as location(s) that specify when and where a conversion formula was applied.

```
<encodingDesc>
  <unitDecl>
    <unitDef  xml:id="keel"  type="weight">
      <label>keel</label>
      <placeName  ref="#england"/>
      <conversion  fromUnit="#chalder"
        toUnit="#keel"  formula="$fromUnit * 20"  from="1421"
        to="1676"/>
      <conversion  fromUnit="#chalder"
        toUnit="#keel"  formula="$fromUnit * 16"  from="1676"
        to="1824"/>
      <desc>Keel was a unit measuring weight of coal. It had been equal to 20 chalders
from 1421 to 1676, and it was made to be equivalent to 16 chalders from 1676 to
1824.</desc>
    </unitDef>
    <unitDef  xml:id="chalder"  type="weight">
      <label>chalder</label>
      <placeName  ref="#england"/>
      <conversion  fromUnit="#bushel"
        toUnit="#chalder"  formula="$fromUnit * 32"  from="1421"
        to="1676"/>
      <conversion  fromUnit="#bushel"
        toUnit="#chalder"  formula="$fromUnit * 36"  from="1676"
        to="1824"/>
      <desc>Chalder was a unit measuring weight of coal. It had been equal to 32
bushels from 1421 to 1676, and it was made to be equivalent to 36 bushels from 1676
to 1824.</desc>
    </unitDef>
    <unitDef  xml:id="bushel"  type="weight">
      <label>bushel</label>
      <placeName  ref="#england"/>
      <desc>Bushel was a unit measuring weight of coal.</desc>
    </unitDef>
  </unitDecl>
</encodingDesc>
```

```
<encodingDesc>
  <unitDecl>
    <unitDef  xml:id="Celsius"
      type="temperature">
      <label>Celsius or Centigrade scale</label>
      <conversion  fromUnit="#Fahrenheit"
        toUnit="#Celsius"  formula="($fromUnit - 32) * (5 div 9)"/>
```

```
     <desc>To convert from the Fahrenheit to the Celsius scale, subtract 32 from the
Celsius temperature and multiply by 5/9.</desc>
   </unitDef>
  </unitDecl>
</encodingDesc>
```

```
<unitDef  type="volume"  xml:id="shunbul">
  <label  xml:lang="ar-Latn-x-ijmes">shunbul</label>
  <label  xml:lang="ar">شنبل</label>
  <conversion  formula="$fromUnit * 2.5"
    fromUnit="#shunbul"  toUnit="#kile"  when="1873"
    where="#acre #tripoli"/>
  <conversion  formula="$fromUnit * 2.25"
    fromUnit="#shunbul"  toUnit="#kile"  when="1878"
    where="#acre #tripoli"/>
  <conversion  formula="$fromUnit * 3"
    from="1891"  fromUnit="#shunbul"  to="1910"
    toUnit="#kile"  where="#aleppo"/>
  <conversion  formula="$fromUnit * 2.5"
    fromUnit="#shunbul"  toUnit="#kile_istanbul"  when="1860"
    where="#aleppo"/>
  <conversion  formula="$fromUnit * 72"
    fromUnit="#shunbul"  toUnit="#madd"  when="1893"
    where="#damascus"/>
  <desc  xml:lang="en">The <foreign  xml:lang="ar-Latn-x-ijmes">shunbul</foreign> was
the main unit to measure grain in the Northern parts of Greater Syria.</desc>
</unitDef>
```

### 2.3.10 The Schema Specification

The <schemaSpec> element contains a schema specification. When this element appears inside <encodingDesc>, it allows embedding of a TEI ODD customization file inside a TEI header; alternatively, this element may be used in the <body> of an ODD document. The use of ODD files, and their relationship to schemas, is described in detail in *23. Documentation Elements*.

A <schemaSpec> element contains all the information needed to generate schemas for a particular TEI customization, and the ODD documentation elements, by reference to the TEI, are more succinct than the schemas derived from them. Therefore you may find it convenient to make a copy of the <schemaSpec> from your project's ODD document inside the <teiHeader> itself, in addition to supplying an external schema and/or ODD file; if the XML file becomes separated from its schema, the schema can be regenerated at any time using the information in the <schemaSpec> element. For example:

```
<encodingDesc>
<!-- Other encoding description elements... -->
  <schemaSpec  ident="myTEICustomization"
    docLang="en"  prefix="tei_"  xml:lang="en"  source="#NONE">
    <moduleRef  key="core"/>
    <moduleRef  key="tei"/>
    <moduleRef  key="header"/>
    <moduleRef  key="textstructure"/>
  </schemaSpec>
</encodingDesc>
```

Alternatively, it is possible to point to an external ODD customization file or schema through use of the <schemaRef> element. More than one <schemaRef> element may be provided for different sources, stages in a workflow, or formats.

```
<schemaRef type="interchangeODD"
  url="http://www.tei-c.org/release/xml/tei/custom/odd/tei_lite.odd"/>
<schemaRef type="interchangeRNG"
  url="http://www.tei-c.org/release/xml/tei/custom/odd/tei_lite.rng"/>
<schemaRef type="projectODD"
  url="file:///schema/project.odd"/>
```

The *url* attribute may be any form of URI pointer including a private URI syntax documented by a <prefixDef>.

### 2.3.11 The Application Information Element

It is sometimes convenient to store information relating to the processing of an encoded resource within its header. Typical uses for such information might be:

- to allow an application to discover that it has previously opened or edited a file, and what version of itself was used to do that;
- to show (through a date) which application last edited the file to allow for diagnosis of any problems that might have been caused by that application;
- to allow users to discover information about an application used to edit the file
- to allow the application to declare an interest in elements of the file which it has edited, so that other applications or human editors may be more wary of making changes to those sections of the file.

The class `model.applicationLike` provides an element, <application>, which may be used to record such information within the <appInfo> element.

`<appInfo>` (application information) records information about an application which has edited the TEI file.

`<application>` provides information about an application which has acted upon the document.

- `@ident` supplies an identifier for the application, independent of its version number or display name.
- `@version` supplies a version number for the application, independent of its identifier or display name.

Each <application> element identifies the current state of one software application with regard to the current file. This element is a member of the `att.datable` class, which provides a variety of attributes for associating this state with a date and time, or a temporal range. The *ident* and *version* attributes should be used to uniquely identify the application and its major version number (for example, ImageMarkupTool 1.5). It is not intended that an application should add a new <application> each time it touches the file.

The following example shows how these elements might be used to document the fact that version 1.5 of an application called 'Image Markup Tool' has an interest in two parts of a document which was last saved on June 6 2006. The parts concerned are accessible at the URLs given as target for the two <ptr> elements.

```
<appInfo>
  <application version="1.5"
    ident="ImageMarkupTool" notAfter="2006-06-01">
    <label>Image Markup Tool</label>
    <ptr target="#P1"/>
    <ptr target="#P2"/>
  </application>
</appInfo>
```

### 2.3.12 Module-Specific Declarations

The elements discussed so far are available to any schema. When the schema in use includes some of the more specialized TEI modules, these make available other more module-specific components of the encoding description.

These are discussed fully in the documentation for the module in question, but are also noted briefly here for convenience.

The <fsdDecl> element is available only when the `iso-fs` module is included in a schema. Its purpose is to document the *feature system declaration* (as defined in chapter *19.11. Feature System Declaration*) underlying any analytic *feature structures* (as defined in chapter *19. Feature Structures*) present in the text documented by this header.

The <metDecl> element is available only when the `verse` module is included in a schema. Its purpose is to document any metrical notation scheme used in the text, as further discussed in section *6.4. Rhyme and Metrical Analysis*. It consists either of a prose description or a series of <metSym> elements.

The <variantEncoding> element is available only when the `textcrit` module is included in a schema. Its purpose is to document the method used to encode textual variants in the text, as discussed in section *13.2. Linking the Apparatus to the Text*.

## 2.4 The Profile Description

The <profileDesc> element is the third major subdivision of the TEI header. It is an optional element, the purpose of which is to enable information characterizing various descriptive aspects of a text or a corpus to be recorded within a single unified framework.

`<profileDesc>` (text-profile description) provides a detailed description of non-bibliographic aspects of a text, specifically the languages and sublanguages used, the situation in which it was produced, the participants and their setting.

In principle, almost any component of the header might be of importance as a means of characterizing a text. The author of a written text, its title or its date of publication, may all be regarded as characterizing it at least as strongly as any of the parameters discussed in this section. The rule of thumb applied has been to exclude from discussion here most of the information which generally forms part of a standard bibliographic style description, if only because such information has already been included elsewhere in the TEI header.

The <profileDesc> element contains elements taken from the `model.profileDescPart` class. The default members of this class are the following :

`<abstract>` contains a summary or formal abstract prefixed to an existing source document by the encoder.

`<creation>` (creation) contains information about the creation of a text.

`<langUsage>` (language usage) describes the languages, sublanguages, registers, dialects, etc. represented within a text.

`<textClass>` (text classification) groups information which describes the nature or topic of a text in terms of a standard classification scheme, thesaurus, etc.

`<correspDesc>` (correspondence description) contains a description of the actions related to one act of correspondence.

`<calendarDesc>` (calendar description) contains a description of the calendar system used in any dating expression found in the text.

These elements are further described in the remainder of this section.

When the `corpus` module described in chapter *16. Language Corpora* is included in a schema, three further elements become available within the <profileDesc> element:

`<textDesc>` (text description) provides a description of a text in terms of its situational parameters.

`<particDesc>` (participation description) describes the identifiable speakers, voices, or other participants in any kind of text or other persons named or otherwise referred to in a text, edition, or metadata.

`<settingDesc>` (setting description) describes the setting or settings within which a language interaction takes place, or other places otherwise referred to in a text, edition, or metadata.

For descriptions of these elements, see section *16.2. Contextual Information*.

When the `transcr` module for the transcription of primary sources described in chapter *12. Representation of Primary Sources* is included in a schema, the following elements become available within the <profileDesc> element:

`<handNotes>` contains one or more <handNote> elements documenting the different hands identified within the source texts.

`<listTranspose>` supplies a list of transpositions, each of which is indicated at some point in a document typically by means of metamarks.

For a description of the <handNotes> element, see section *12.3.2.1. Document Hands*. Its purpose is to group together a number of <handNote> elements, each of which describes a different hand or equivalent identified within a manuscript. The <handNote> element can also appear within a structured manuscript description, when the `msdescription` module described in chapter *11. Manuscript Description* is included in a schema. For this reason, the <handNote> element is actually declared within the header module, but is only accessible to a schema when one or other of the `transcr` or `msdescription` modules is included in a schema. See further the discussion at *12.3.2.1. Document Hands*.

The <listTranspose> element is discussed in detail in section *12.3.4.5. Transpositions*.

### 2.4.1 Creation

The <creation> element contains phrases describing the origin of the text, e.g. the date and place of its composition.

`<creation>` (creation) contains information about the creation of a text.

The date and place of composition are often of particular importance for studies of linguistic variation; since such information cannot be inferred with confidence from the bibliographic description of the copy text, the <creation> element may be used to provide a consistent location for this information:

```
<creation>
  <date when="1992-08">August 1992</date>
  <rs type="city">Taos, New Mexico</rs>
</creation>
```

### 2.4.2 Language Usage

The <langUsage> element is used within the <profileDesc> element to describe the languages, sublanguages, registers, dialects, etc. represented within a text. It contains one or more <language> elements, each of which provides information about a single language, notably the quantity of that language present in the text. Note that this element should *not* be used to supply information about any non-standard characters or glyphs used by this language; such information should be recorded in the <charDecl> element in the encoding description (see further *5. Characters, Glyphs, and Writing Modes*).

`<langUsage>` (language usage) describes the languages, sublanguages, registers, dialects, etc. represented within a text.

`<language>` (language) characterizes a single language or sublanguage used within a text.

`@usage` specifies the approximate percentage of the text which uses this language.

`@ident` (identifier) Supplies a language code constructed as defined in BCP 47 which is used to identify the language documented by this element, and which may be referenced by the global *xml:lang* attribute.

A <language> element may be supplied for each different language used in a document. If used, its *ident* attribute should specify an appropriate language identifier, as further discussed in section *vi.1 Language Identification*. This is particularly important if extended language identifiers have been used as the value of *xml:lang* attributes elsewhere in the document.

Here is an example of the use of this element:

```
<langUsage>
  <language ident="fr-CA" usage="60">Québecois</language>
  <language ident="en-CA" usage="20">Canadian business English</language>
  <language ident="en-GB" usage="20">British English</language>
</langUsage>
```

### 2.4.3 The Text Classification

The <textClass> element is used to classify a text in some way.

`<textClass>` (text classification) groups information which describes the nature or topic of a text in terms of a standard classification scheme, thesaurus, etc.

Text classification may be carried out according to one or more of the following methods:

- by reference to a recognized international classification such as the Dewey Decimal Classification, the Universal Decimal Classification, the Colon Classification, the Library of Congress Classification, or any other system widely used in library and documentation work
- by providing a set of keywords, as provided for example by British Library or Library of Congress Cataloguing in Publication data
- by referencing any other taxonomy of text categories recognized in the field concerned, or peculiar to the material in hand; this may include one based on recurring sets of values for the situational parameters defined in section *16.2.1. The Text Description*, or the demographic elements described in section *16.2.2. The Participant Description*

The last of these may be particularly important for dealing with existing corpora or collections, both as a means of avoiding the expense or inconvenience of reclassification and as a means of documenting the organizing principles of such materials.

The following elements are provided for this purpose:

`<keywords>` (keywords) contains a list of keywords or phrases identifying the topic or nature of a text.

`@scheme` identifies the controlled vocabulary within which the set of keywords concerned is defined, for example by a <taxonomy> element, or by some other resource.

`<classCode>` (classification code) contains the classification code used for this text in some standard classification system.

`@scheme` identifies the classification system in use, as defined by, e.g. a <taxonomy> element, or some other resource.

`<catRef>` (category reference) specifies one or more defined categories within some taxonomy or text typology.

The <keywords> element simply categorizes an individual text by supplying a list of keywords which may describe its topic or subject matter, its form, date, etc. In some schemes, the order of items in the list is significant, for example, from major topic to minor; in others, the list has an organized substructure of its own. No recommendations are made here as to which method is to be preferred. Wherever possible, such keywords should be taken from a recognized source, such as the British Library/Library of Congress Cataloguing in Publication data in the case of printed books, or a published thesaurus appropriate to the field.

The *scheme* attribute is used to indicate the source of the keywords used, in the case where such a source exists. If the keywords are taken from an externally defined authority which is available online, this attribute should point directly to it, as in the following examples:

```
<keywords  scheme="http://classificationweb.net">
  <term>Babbage, Charles</term>
  <term>Mathematicians - Great Britain - Biography</term>
</keywords>
```

```
<keywords  scheme="http://id.loc.gov/authorities/about.html#lcsh">
  <term>English literature -- History and criticism -- Data processing.</term>
  <term>English literature -- History and criticism -- Theory, etc.</term>
  <term>English language -- Style -- Data processing.</term>
  <term>Style, Literary -- Data processing.</term>
</keywords>
```

If the authority file is not available online, but is generally recognized and commonly cited, a bibliographic description for it should be supplied within the <taxonomy> element described in section *2.3.7. The Classification Declaration*; the *scheme* attribute may then reference that <taxonomy> element by means of its identifier in the usual way:

```
<keywords  scheme="#welch">
  <term>ceremonials</term>
  <term>fairs</term>
  <term>street life</term>
</keywords>
<!-- elsewhere in the document -->
<taxonomy  xml:id="welch">
  <bibl>
    <title>Notes on London Municipal Literature, and a Suggested
        Scheme for Its Classification</title>
    <author>Charles Welch</author>
    <edition>1895</edition>
  </bibl>
</taxonomy>
```

If no authority file exists, perhaps because the keywords used were assigned directly by an author, the *scheme* attribute should be omitted.

Alternatively, if the keyword vocabulary itself is locally defined, the *scheme* attribute will point to the local definition, which will typically be held in a <taxonomy> element within the <classDecl> part of the encoding description (see section *2.3.7. The Classification Declaration*).

The <classCode> element also categorizes an individual text, by supplying a numerical or other code rather than descriptive terms. Such codes constitute a recognized classification scheme, such as the Dewey Decimal Classification. On this element, the *scheme* attribute is required; it indicates the source of the classification scheme in the same way as for keywords: this may be a pointer of any kind, either to a TEI element, possibly in the current document, as in the <keywords> examples above, or to some canonical source for the scheme, as in the following example:

```
<classCode  scheme="http://www.udcc.org/udcsummary/php/index.php">005.756</classCode>
```

The <catRef> element categorizes an individual text by pointing to one or more <category> elements using the *target* attribute, which it inherits from the `att.pointing` class. The <category> element (which is fully described in section *2.3.7. The Classification Declaration*) holds information about a particular classification or category within a given taxonomy. Each such category must have a unique identifier, which may be supplied as the value of the *target* attribute for <catRef> elements which are regarded as falling within the category indicated.

A text may, of course, fall into more than one category, in which case more than one identifier may be supplied as the value for the *target* attribute on the <catRef> element, as in the following example:

```
<catRef  target="#b.a4 #b.d2"/>
```

The *scheme* attribute may be supplied to specify the taxonomy to which the categories identified by the target attribute belong, if this is not adequately conveyed by the resource pointed to. For example,

```
<catRef  target="#b.a4 #b.d2"
  scheme="http://www.example.com/browncorpus"/>
<catRef  target="http://www.example.com/SUC/#A45"/>
```

Here the same text has been classified as of categories b.a4 and b.d2 within the Brown classification scheme (presumed to be available from `http://www.example.com/browncorpus`), and as of category 'A45' within the SUC classification scheme documented at the URL given.

In general, it is a matter of style whether to use a single <catRef> with multiple identifiers in the value of *target* or multiple <catRef> elements, each with a single identifier in the value of *target*. However, note that maintenance of a TEI document with a large number of values within a single *target* can be cumbersome.

The distinction between the <catRef> and <classCode> elements is that the values used as identifying codes are exhaustively enumerated for the former, typically within the TEI header. In the latter case, however, the values use any externally-defined scheme, and therefore may be taken from a more open-ended descriptive classification system.

### 2.4.4 Abstracts

The main purpose of the <abstract> element is to supply a brief resume or abstract for an article which was originally published without such a component. An abstract or summary forming part of the document at its creation should usually appear in the front matter (<front>) of the document.

```
<profileDesc>
  <abstract>
    <p>This paper is a draft studying
        various aspects of using the TEI
        as a reference serialization framework
        for LMF. Comments are welcome to bring
        this to a useful document for the
        community.
    </p>
  </abstract>
</profileDesc>
```

Abstracts may be provided in multiple languages, distinguished by the *xml:lang* attribute:

```
<profileDesc>
  <abstract  xml:lang="en">
    <p>The recent archaeological emphasis
        on the study of settlement patterns,
        landscape and palaeoenvironments has
        shaped and re-shaped our understanding
        of the Viking settlement of Iceland.
        This paper reviews the developments
        in Icelandic archaeology, examining
        both theoretical and practical advances.
        Particular attention is paid to new
        ideas in terms of settlement patterns
        and resource exploitation. Finally,
        some of the key studies of the ecological
        consequences of the Norse
    <foreign  xml:lang="is">landnám</foreign>
        are presented. </p>
  </abstract>
  <abstract  xml:lang="fr">
    <p>L'accent récent des
        recherches archéologiques sur l'étude des
        configurations spatiales des colonies, de la
        géographie des sites ainsi que des éléments
        paléo-environnementaux nous mène à réexaminer
        et réévaluer nos connaissances acquises sur
        la colonisation de l'Islande par les Vikings.
        Cet article passe en revue le développement
        de l'archéologie islandaise en examinant les
        progrès théoriques et pratiques en la matière.
        Une attention particulière est portée sur
        l'étude des configurations spatiales des
```

```
        colonies ainsi qu'une considération des
        questions d'exploitation des ressources.
        Finalement, l'article présente un aperçu des
        études principales qui traitent des
        conséquences écologiques du
    <foreign  xml:lang="is">landnám</foreign>
        islandais.</p>
  </abstract>
</profileDesc>
```

The same element may be used to provide other summary information supplied by the encoder, perhaps grouped together into a list of discrete items:

```
<profileDesc>
  <abstract>
    <list>
      <item>An annual HBC supply ship is
            set to the North West Coast for mid-September.</item>
      <item>
        <name  key="pelly_jh">Pelly</name> writes
            to ascertain the British Government's plans
            for the lands associated with the Oregon Treaty;
            he wants to know what will happen to the HBC's
            establishment on the southern <name  type="place"
         key="vancouver_island">Vancouver Island</name>.
            He adds that a former Crown grant, an 1838 exclusive
            trade-grant for the lands in question, has yet to
            expire.</item>
      <item>The minutes discuss the nature of the HBC's
            original entitlements and question whether or not,
            and in what capacity, the Oregon Treaty affects the
            HBC's position. The majority council further
            investigation, and to reply cautiously and
            judiciously to the HBC inquiry.</item>
      <item>A
            summary of a meeting with <name  key="pelly_jh">Pelly</name> is offered in
            order to elucidate the HBC's intentions.</item>
      <item>
        <name  key="grey_hg">Lord Grey</name> calls
            for greater consideration on the issue of
            colonization; he asks that <name  key="stephen_j">Stephen</name> write the
Company,
            asking them to detail their intentions, and to
            state their legal opinion for entitlement.
      </item>
    </list>
  </abstract>
</profileDesc>
```

### 2.4.5 Calendar Description

The <calendarDesc> element is used within the <profileDesc> element to document objects referenced by means of either the *calendar* attribute on <date> or the *datingMethod* attribute on any member of the `att.datable` class.

`<calendarDesc>` (calendar description) contains a description of the calendar system used in any dating expression found in the text.

This element may contain one or more <calendar> elements:

`<calendar>` (calendar) describes a calendar or dating system used in a dating formula in the text.

Each such element contains one or more paragraphs of description for the calendar system concerned, and also supplies an identifying code for it as the value of its *xml:id* attribute. A <calendar> element may also have a *type* (and *subtype*) attribute in order to categorize the calendar in any convenient way. In the following example, *type* is used to differentiate real calendar systems in actual use from those used in a fictional environment.

```
<calendarDesc>
  <calendar xml:id="Gregorian" type="actual">
    <p>Gregorian calendar</p>
  </calendar>
  <calendar xml:id="Stardate"
    type="fictional">
    <p>Fictional Stardate (from Star Trek series)</p>
  </calendar>
  <calendar xml:id="BP" type="actual">
    <p>Calendar years before present (measured from 1950)</p>
  </calendar>
</calendarDesc>
```

This identifying code may then be referenced from any element supplying a date expressed using that calendar system:

```
<p>Captain's log <date calendar="#stardate">stardate 23.9 rounded off
    to the nearest decimal point</date>...</p>
```

See *14.4.4. Using Non-Gregorian Calendars* for details of the usage of dating attributes in conjunction with <calendar> elements in the <teiHeader>.

### 2.4.6 Correspondence Description

The <correspDesc> element is used within the <profileDesc> element to provide detailed correspondence-specific metadata, concerning in particular the communicative aspects (sending, receiving, forwarding etc.) associated with an act of correspondence.

This information is complementary to the detailed descriptions of physical objects (such as letters) associated with correspondence acts, which are typically provided by the <sourceDesc> element.

`<correspDesc>` (correspondence description) contains a description of the actions related to one act of correspondence.

The <correspDesc> element contains the elements <correspAction> and <correspContext>, describing the actions identified and the context in which the correspondence occurs respectively.

`<correspAction>` (correspondence action) contains a structured description of the place, the name of a person/organization and the date related to the sending/receiving of a message or any other action related to the correspondence.

`@type` describes the nature of the action.

`<correspContext>` (correspondence context) provides references to preceding or following correspondence related to this piece of correspondence.

Acts of correspondence typically do not occur in isolation from each other. The <correspContext> element is used to group references relevant to the item of correspondence being described, typically to other items such as the item to which it is a reply, or the item which replies to it:

```
<correspContext>
  <ref type="prev" target="#CLF0102">Previous letter of <persName>Chamisso</persName>
to <persName>de La
```

```
        Foye</persName>: <date  when="1807-01-16">16 January 1807</date>
  </ref>
  <ref  type="next"  target="#CLF0104">Next letter of <persName>Chamisso</persName> to
<persName>de La Foye</persName>:
  <date  when="1810-05-07">07 May 1810</date>
  </ref>
</correspContext>
```

Many types of correspondence actions may be distinguished. The *type* attribute should be used to indicate the type of action being documented, using values such as those suggested above.

The following simple example uses <correspAction> to describe the sending of a letter by Adelbert von Chamisso from Vertus on 29 January 1807 to Louis de La Foye at Caen. The date of reception is unknown:

```
<correspAction  type="sent">
  <persName>Adelbert von Chamisso</persName>
  <placeName>Vertus</placeName>
  <date  when="1807-01-29"/>
</correspAction>
<correspAction  type="received">
  <persName>Louis de La Foye</persName>
  <placeName>Caen</placeName>
  <date>unknown</date>
</correspAction>
```

Note the use of the *when* attribute described in *3.6.4. Dates and Times* to provide a normalized form of the date. The content of the <date> element may be omitted, since no underlying source is being transcribed.

Several senders, recipients, etc. can be specified for a single <correspAction> if the action is considered to apply to them all acting as a single group. In the following example, two people are considered to have received the piece of correspondence.

```
<correspAction  type="received">
  <persName>Hermann Hesse</persName>
  <persName>Ninon Hesse</persName>
  <placeName>Montagnola</placeName>
</correspAction>
```

The *subtype* attribute may be used to further distinguish types of action. In the following example, an e-mail is sent to two people directly and to a third by 'carbon copy' (*cc*).

```
<correspAction  type="sent">
  <persName>PN0001</persName>
  <date  when="1999-06-02"/>
</correspAction>
<correspAction  type="received"  subtype="to">
  <persName>PN0002</persName>
</correspAction>
<correspAction  type="received"  subtype="to">
  <persName>PN0003</persName>
</correspAction>
<correspAction  type="received"  subtype="cc">
  <persName>PN0004</persName>
</correspAction>
```

The same person may be associated with several actions. For example, it will often be the case that the author and sender of a message are identical, and that many individual letters will need to be associated with the same person. The *sameAs* attribute defined in section *17.6. Identical Elements and Virtual Copies* may be used to indicate that the same name applies to many actions. Its value will usually be the identifier of an element defining the person or name concerned, which is supplied elsewhere in the document.

```
<correspAction type="sent">
  <name sameAs="#author"/>
</correspAction>
```

It is assumed that each correspondence action applies to a single act of communication. However, it may be the case that the same physical object is involved in several such acts, if for example person A sends a letter to person B, who then annotates it and sends it on to person C, or if persons A and B both use the same document to convey quite different messages. In such situations, multiple <correspDesc> elements should be supplied, one for each communication. In the following example, the same document contains distinct messages, sent by two different people to the same destination: <correspDesc> is used for each separately:

```
<correspDesc xml:id="message1">
  <correspAction type="sent">
    <persName>John Gneisenau Neihardt</persName>
    <placeName>Branson (Montgomery)</placeName>
    <date when="1932-12-17"/>
  </correspAction>
  <correspAction type="received">
    <persName xml:id="JTH">Julius Temple House</persName>
    <placeName>New York</placeName>
  </correspAction>
</correspDesc>
<correspDesc xml:id="message2">
  <correspAction type="sent">
    <persName>Enid Neihardt</persName>
    <placeName>Branson (Montgomery)</placeName>
    <date when="1932-12-17"/>
  </correspAction>
  <correspAction type="received">
    <persName sameAs="#JTH"/>
    <placeName>New York</placeName>
  </correspAction>
</correspDesc>
```

## 2.5 Non-TEI Metadata

Projects often maintain metadata about their TEI documents in more than one form or system. For example, a project may have a database of bibliographic information on the set of documents they intend to encode. From this database, both a MARC record and a <teiHeader> are generated. The document is then encoded, during which process additional information is added to the <teiHeader> manually. Then, when the document is published on the web, a Dublin Core record is generated for discoverability of the resource. It is sometimes advantageous to store some or all of the non-TEI metadata in the TEI file.

Such non-TEI data may be placed anywhere within a TEI file (other than as the root element), as it does not affect TEI conformance. However, it is easier for humans to manage these kinds of data if they are grouped together in a single location. In addition, such grouping makes it easy to avoid accidentally flagging non-TEI data as errors during validation of the file against a TEI schema. The <xenoData> element, which may appear in the TEI header after the <fileDesc> but before the optional <revisionDesc>, is provided for this purpose. Where the use of <xenoData> is primarily to store linked data, contextual information, or stand-off annotations that refer to the contents of the TEI file itself, the <xenoData> element may be placed inside the <standOff> element instead.

`<xenoData>` (non-TEI metadata) provides a container element into which metadata in non-TEI formats may be placed.

The <xenoData> element may contain anything except TEI elements. It may contain one or more elements from outside the TEI[42] or data in some non-XML text format.[43]

In the following example, the prefix `rdf` is bound to the namespace `http://www.w3.org/1999/02/22-rdf-syntax-n` the prefix `dc` is bound to the namespace `http://purl.org/dc/elements/1.1/`, and the prefix `cc` is bound to the namespace `http://web.resource.org/cc/`.

```
<xenoData
   xmlns:cc="http://web.resource.org/cc/"
   xmlns:dc="http://purl.org/dc/elements/1.1/"
   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:RDF>
    <cc:Work  rdf:about="">
      <dc:title>Applied Software Project Management - review</dc:title>
      <dc:type  rdf:resource="http://purl.org/dc/dcmitype/Text"/>
      <dc:license  rdf:resource="http://creativecommons.org/licenses/by-sa/2.0/uk/"/>
    </cc:Work>
    <cc:License  rdf:about="http://creativecommons.org/licenses/by-sa/2.0/uk/">
      <cc:permits  rdf:resource="http://web.resource.org/cc/Reproduction"/>
      <cc:permits  rdf:resource="http://web.resource.org/cc/Distribution"/>
      <cc:requires  rdf:resource="http://web.resource.org/cc/Notice"/>
      <cc:requires  rdf:resource="http://web.resource.org/cc/Attribution"/>
      <cc:permits  rdf:resource="http://web.resource.org/cc/DerivativeWorks"/>
      <cc:requires  rdf:resource="http://web.resource.org/cc/ShareAlike"/>
    </cc:License>
  </rdf:RDF>
</xenoData>
```

## 2.6 The Revision Description

The final sub-element of the TEI header, the <revisionDesc> element, provides a detailed change log in which each change made to a text may be recorded. Its use is optional but highly recommended. It provides essential information for the administration of large numbers of files which are being updated, corrected, or otherwise modified as well as extremely useful documentation for files being passed from researcher to researcher or system to system. Without change logs, it is easy to confuse different versions of a file, or to remain unaware of small but important changes made in the file by some earlier link in the chain of distribution. No significant change should be made in any TEI-conformant file without corresponding entries being made in the change log.

`<revisionDesc>` (revision description) summarizes the revision history for a file.

`<listChange>` groups a number of change descriptions associated with either the creation of a source text or the revision of an encoded text.

`<change>` (change) documents a change or set of changes made during the production of a source document, or during the revision of an electronic file.

The main purpose of the revision description is to record changes in the text to which a header is prefixed. However, it is recommended TEI practice to include entries also for significant changes in the header itself (other than the revision description itself, of course). At the very least, an entry should be supplied indicating the date of creation of the header.

The log consists of a list of entries, one for each change. Changes may be grouped and organised using either the <listChange> element described in section *12.7. Identifying Changes and Revisions* or the simple <list> element described in section *3.8. Lists*. Alternatively, a simple sequence of <change> elements may be given. The attributes

[42] As is always the case when mixing elements from different namespaces in an XML document, the namespace of these non-TEI elements must be declared either on the elements themselves or on an ancestor element.

[43] As is always the case when using text inside an XML document, certain characters cannot occur in their normal form, and must be 'escaped'. The most common of these are LESS-THAN SIGN ('<', U+003C) and AMPERSAND ('&', U+0026), which may be escaped with `<` and `&` respectively. See *v.7.1 Character References*.

*when* and *who* may be supplied for each <change> element to indicate its date and the person responsible for it respectively. The description of the change itself can range from a simple phrase to a series of paragraphs. If a number is to be associated with one or more changes (for example, a revision number), the global *n* attribute may be used to indicate it.

It is recommended to give changes in reverse chronological order, most recent first.

For example:

```
<!-- ... --><revisionDesc>
  <change  n="RCS:1.39"  when="2007-08-08"
    who="#jwernimo.lrv">Changed <val>drama.verse</val>
    <gi>lg</gi>s to <gi>p</gi>s. <note>we have opened a discussion about the need for a
new
        value for <att>type</att> of <gi>lg</gi>, <val>drama.free.verse</val>, in order
to address
        the verse of Behn which is not in regular iambic pentameter. For the time being
these
        instances are marked with a comment note until we are able to fully consider
the best way
        to encode these instances.</note>
  </change>
  <change  n="RCS:1.33"  when="2007-06-28"
    who="#pcaton.xzc">Added <att>key</att> and <att>reg</att>
     to <gi>name</gi>s.</change>
  <change  n="RCS:1.31"  when="2006-12-04"
    who="#wgui.ner">Completed renovation. Validated.</change>
</revisionDesc>
```

In the above example, the *who* attributes point to <respStmt> elements which have been included earlier in the <titleStmt> of the same header:

```
<titleStmt>
  <title>The Amorous Prince, or, the Curious Husband, 1671</title>
  <author>
    <persName  ref="#abehn.aeh">Behn, Aphra</persName>
  </author>
  <respStmt  xml:id="pcaton.xzc">
    <persName>Caton, Paul</persName>
    <resp>electronic publication editor</resp>
  </respStmt>
  <respStmt  xml:id="wgui.ner">
    <persName>Gui, Weihsin</persName>
    <resp>encoder</resp>
  </respStmt>
  <respStmt  xml:id="jwernimo.lrv">
    <persName>Wernimont, Jacqueline</persName>
    <resp>encoder</resp>
  </respStmt>
</titleStmt>
```

There is however no requirement that the <respStmt> be used for this person, or that the elements indicated be contained within the same document. A project might for example maintain a separate document listing all of its personnel in which they were represented using the <person> element described in *16.2.2. The Participant Description*.

## 2.7 Minimal and Recommended Headers

The TEI header allows for the provision of a very large amount of information concerning the text itself, its source, its encodings, and revisions of it, as well as a wealth of descriptive information such as the languages it uses and the

situation(s) in which it was produced, together with the setting and identity of participants within it. This diversity and richness reflects the diversity of uses to which it is envisaged that electronic texts conforming to these Guidelines will be put. It is emphatically *not* intended that all of the elements described above should be present in every TEI header.

The amount of encoding in a header will depend both on the nature and the intended use of the text. At one extreme, an encoder may expect that the header will be needed only to provide a bibliographic identification of the text adequate to local needs. At the other, wishing to ensure that their texts can be used for the widest range of applications, encoders will want to document as explicitly as possible both bibliographic and descriptive information, in such a way that no prior or ancillary knowledge about the text is needed in order to process it. The header in such a case will be very full, approximating to the kind of documentation often supplied in the form of a manual. Most texts will lie somewhere between these extremes; textual corpora in particular will tend more to the latter extreme. In the remainder of this section we demonstrate first the minimal, and next a commonly recommended, level of encoding for the bibliographic information held by the TEI header.

Supplying only the minimal level of encoding required, the TEI header of a single text might look like the following example:

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>Thomas Paine: Common sense, a
            machine-readable transcript</title>
      <respStmt>
        <resp>compiled by</resp>
        <name>Jon K Adams</name>
      </respStmt>
    </titleStmt>
    <publicationStmt>
      <distributor>Oxford Text Archive</distributor>
    </publicationStmt>
    <sourceDesc>
      <bibl>The complete writings of Thomas Paine, collected and edited
            by Phillip S. Foner (New York, Citadel Press, 1945)</bibl>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

The only mandatory component of the TEI header is the <fileDesc> element. Within this, <titleStmt>, <publicationStmt>, and <sourceDesc> are all required constituents. Within the title statement, a title is required, and an author should be specified, even if it is *unknown*, as should some additional statement of responsibility, here given by the <respStmt> element. Within the <publicationStmt>, a publisher, distributor, or other agency responsible for the file must be specified. Finally, the source description should contain at the least a loosely structured bibliographic citation identifying the source of the electronic text if (as is usually the case) there is one.

We now present the same example header, expanded to include additionally recommended information, adequate to most bibliographic purposes, in particular to allow for the creation of an AACR2-conformant bibliographic record. We have also added information about the encoding principles used in this (imaginary) encoding, about the text itself (in the form of Library of Congress subject headings), and about the revision of the file.

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>Common sense, a machine-readable transcript</title>
      <author>Paine, Thomas (1737-1809)</author>
      <respStmt>
        <resp>compiled by</resp>
```

```
      <name>Jon K Adams</name>
    </respStmt>
  </titleStmt>
  <editionStmt>
    <edition>
      <date>1986</date>
    </edition>
  </editionStmt>
  <publicationStmt>
    <distributor>Oxford Text Archive.</distributor>
    <address>
      <addrLine>Oxford University Computing Services,</addrLine>
      <addrLine>13 Banbury Road,</addrLine>
      <addrLine>Oxford OX2 6RB,</addrLine>
      <addrLine>UK</addrLine>
    </address>
  </publicationStmt>
  <notesStmt>
    <note>Brief notes on the text are in a
          supplementary file.</note>
  </notesStmt>
  <sourceDesc>
    <biblStruct>
      <monogr>
        <editor>Foner, Philip S.</editor>
        <title>The collected writings of Thomas Paine</title>
        <imprint>
          <pubPlace>New York</pubPlace>
          <publisher>Citadel Press</publisher>
          <date>1945</date>
        </imprint>
      </monogr>
    </biblStruct>
  </sourceDesc>
</fileDesc>
<encodingDesc>
  <samplingDecl>
    <p>Editorial notes in the Foner edition have not
          been reproduced. </p>
    <p>Blank lines and multiple blank spaces, including paragraph
          indents, have not been preserved. </p>
  </samplingDecl>
  <editorialDecl>
    <correction  status="high"
      method="silent">
      <p>The following errors
             in the Foner edition have been corrected:
      <list>
         <item>p. 13 l. 7 cotemporaries contemporaries</item>
         <item>p. 28 l. 26 [comma] [period]</item>
         <item>p. 84 l. 4 kin kind</item>
         <item>p. 95 l. 1 stuggle struggle</item>
         <item>p. 101 l. 4 certainy certainty</item>
         <item>p. 167 l. 6 than that</item>
         <item>p. 209 l. 24 publshed published</item>
       </list>
      </p>
```

```
      </correction>
      <normalization>
        <p>No normalization beyond that performed
                by Foner, if any. </p>
      </normalization>
      <quotation  marks="all">
        <p>All double quotation marks
                rendered with ", all single quotation marks with
                apostrophe. </p>
      </quotation>
      <hyphenation  eol="none">
        <p>Hyphenated words that appear at the
                end of the line in the Foner edition have been reformed.</p>
      </hyphenation>
      <stdVals>
        <p>The values of <att>when-iso</att> on the <gi>time</gi>
                element always end in the format <val>HH:MM</val> or
        <val>HH</val>; i.e., seconds, fractions thereof, and time
                zone designators are not present.</p>
      </stdVals>
      <interpretation>
        <p>Compound proper names are marked. </p>
        <p>Dates are marked. </p>
        <p>Italics are recorded without interpretation. </p>
      </interpretation>
    </editorialDecl>
    <classDecl>
      <taxonomy  xml:id="lcsh">
        <bibl>Library of Congress Subject Headings</bibl>
      </taxonomy>
      <taxonomy  xml:id="lc">
        <bibl>Library of Congress Classification</bibl>
      </taxonomy>
    </classDecl>
  </encodingDesc>
  <profileDesc>
    <creation>
      <date>1774</date>
    </creation>
    <langUsage>
      <language  ident="en"  usage="100">English.</language>
    </langUsage>
    <textClass>
      <keywords  scheme="#lcsh">
        <term>Political science</term>
        <term>United States — Politics and government —
                Revolution, 1775-1783</term>
      </keywords>
      <classCode  scheme="#lc">JC 177</classCode>
    </textClass>
  </profileDesc>
  <revisionDesc>
    <change  when="1996-01-22"  who="#MSM"> finished proofreading </change>
    <change  when="1995-10-30"  who="#LB"> finished proofreading </change>
    <change  notBefore="1995-07-04"  who="#RG"> finished data entry at end of term
</change>
    <change  notAfter="1995-01-01"  who="#RG"> began data entry before New Year 1995
```

```
</change>
  </revisionDesc>
</teiHeader>
```

Many other examples of recommended usage for the elements discussed in this chapter are provided here, in the reference index and in the associated tutorials.

## 2.8 Note for Library Cataloguers

A strong motivation in preparing the material in this chapter was to provide in the TEI header a viable chief source of information for cataloguing computer files. The TEI header is not a library catalogue record, and so will not make all of the distinctions essential in standard library work. It also includes much information generally excluded from standard bibliographic descriptions. It is the intention of the developers, however, to ensure that the information required for a catalogue record be retrievable from the TEI file header, and moreover that the mapping from the one to the other be as simple and straightforward as possible. Where the correspondence is not obvious, it may prove useful to consult one of the works which were influential in developing the content of the TEI header. These include:

*ISBD: International Standard Bibliographic Description* is an international standard setting out what information should be recorded in a description of a bibliographical item. Until a consolidated edition published in 2011, there was a general standard called ISBD(G) and separate ISBDs covering different types of material, e.g. ISBD(M) for monographs, ISBD(ER) for electronic resources. These separate ISBDs follow the same general scheme as the main ISBD(G), but provide appropriate interpretations for the specific materials under consideration.

The *Anglo-American Cataloguing Rules (second edition)* were published in 1978, with revisions appearing periodically through 2005. AACR2 provides guidelines for the construction of catalogues in general libraries in the English-speaking world. AACR2 is explicitly based on the general framework of the ISBD(G) and the subsidiary ISBDs: it gives a description of how to describe bibliographic items and how to create access points such as subject or name headings and uniform titles. Other national cataloguing codes exist as well, including the Z44 series of standards from issued by the Association française de normalisation (AFNOR), *Regeln für die alphabetische Katalogisierung in wissenschaftlichen Bibliotheken* (RAK-WB), *Regole italiane di catalogazione per autore* (RICA), and *Система стандартов по информации, библиотечному и издательскому делу. Библиографическая запись. Библиографическое описание. Общие требования и правила составления* (ГОСТ 7.1).

The *American National Standard for Bibliographic References* was an American national standard governing bibliographic references for use in bibliographies, end-of-work lists, references in abstracting and indexing publications, and outputs from computerized bibliographic data bases. A revised version is maintained by the National Information Standards Organization (NISO). The related ISO standard is ISO 690. Other relevant national standards include BS 5605:1990, BS 6371:1983, DIN 1505-2, and ГОСТ 7.0.5.

Since the TEI file description elements are based on the ISBD areas, it should be possible to use the content of file description as the basis for a catalog record for a TEI document. However, cataloguers should be aware that the permissive nature of the TEI Guidelines may lead to divergences between practice in using the TEI file description and the comparatively strict recommendations of AACR2 and other national cataloguing codes. Such divergences as the following may preclude automatic generation of catalogue records from TEI headers:

- The TEI Guidelines do not require that text be transcribed from the 'chief source of information' using normalized capitalization and punctuation .
- The TEI title statement may not categorize constituent titles in the same way as prescribed by a national cataloguing code.
- The TEI title statement contains authors, editors, and other responsible parties in separate elements, with names which may not have been normalized; it does not necessarily contain a single statement of responsibility .
- There is no specific place in a TEI header to specify the *main entry* or *added entries* ( *name or title headings under which a catalogue record is filed*) for the catalogue record.

- The TEI header does not require use of a particular vocabulary for subject headings nor require the use of subject headings.

## 2.9 The TEI Header Module

The module described in this chapter makes available the following components:

**Module header**: **The TEI header**

- *Elements defined*: abstract appInfo application authority availability biblFull cRefPattern calendar calendarDesc catDesc catRef category change citeData citeStructure classCode classDecl conversion correction correspAction correspContext correspDesc creation distributor edition editionStmt editorialDecl encodingDesc extent fileDesc funder geoDecl handNote hyphenation idno interpretation keywords langUsage language licence listChange listPrefixDef namespace normalization notesStmt prefixDef principal profileDesc projectDesc publicationStmt punctuation quotation refState refsDecl rendition revisionDesc samplingDecl schemaRef scriptNote segmentation seriesStmt sourceDesc sponsor stdVals styleDefDecl tagUsage tagsDecl taxonomy teiHeader textClass titleStmt unitDecl unitDef xenoData
- *Classes defined*: att.citeStructurePart att.patternReplacement

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 3 Elements Available in All TEI Documents

This chapter describes elements which may appear in any kind of text and the tags used to mark them in all TEI documents. Most of these elements are freely floating phrases, which can appear at any point within the textual structure, although they should generally be contained by a higher-level element of some kind (such as a paragraph). A few of the elements described in this chapter (for example, bibliographic citations and lists) have a comparatively well-defined internal structure, but most of them have no consistent inner structure of their own. In the general case, they contain only a few words, and are often identifiable in a conventionally printed text by the use of typographic conventions such as shifts of font, use of quotation or other punctuation marks, or other changes in layout.

This chapter begins by describing the <p> tag used to mark paragraphs, the prototypical formal unit for running text in many TEI modules. This is followed, in section *3.2. Treatment of Punctuation*, by a discussion of some specific problems associated with the interpretation of conventional punctuation, and the methods proposed by these Guidelines for resolving ambiguities therein.

The next section (section *3.3. Highlighting and Quotation*) describes a number of phrase-level elements commonly marked by typographic features (and thus well-represented in conventional markup languages). These include features commonly marked by font shifts (section *3.3.2. Emphasis, Foreign Words, and Unusual Language*) and features commonly marked by quotation marks (section *3.3.3. Quotation*) as well as such features as terms, cited words, and glosses (section *3.4. Terms and Glosses, Ruby Annotations, and Equivalents and Descriptions*).

Section *3.5. Simple Editorial Changes* introduces some phrase-level elements which may be used to record simple editorial interventions, such as emendation or correction of the encoded text. The elements described here constitute a simple subset of the full mechanisms for encoding such information (described in full in chapter *12. Representation of Primary Sources*), which should be adequate to most commonly encountered situations.

The next section (section *3.6. Names, Numbers, Dates, Abbreviations, and Addresses*) describes several phrase-level and inter-level elements which, although often of interest for analysis or processing, are rarely explicitly identified in conventional printing. These include names (section *3.6.1. Referring Strings*), numbers and measures (section *3.6.3. Numbers and Measures*), dates and times (section *3.6.4. Dates and Times*), abbreviations (section *3.6.5. Abbreviations and Their Expansions*), and addresses (section *3.6.2. Addresses*).

In the same way, the following section (section *3.7. Simple Links and Cross-References*) presents only a subset of the facilities available for the encoding of cross-references or text-linkage. The full story may be found in chapter *17. Linking, Segmentation, and Alignment*; the tags presented here are intended to be usable for a wide variety of simple applications.

Sections *3.8. Lists*, and *3.9. Notes, Annotation, and Indexing*, describe two kinds of quasi-structural elements: lists and notes. These may appear either within chunk-level elements such as paragraphs, or between them. Several kinds of lists are catered for, of an arbitrary complexity. The section on notes discusses both notes found in the source and simple mechanisms for adding annotations of an interpretive nature during the encoding; again, only a subset of the facilities described in full elsewhere (specifically, in chapter *18. Simple Analytic Mechanisms*) is discussed.

Section *3.10. Graphics and Other Non-textual Components* introduces some simple ways of representing graphic or other non-textual content found in a text. A fuller discussion of the multimedia facilities supported by these Guidelines may be found in chapters *15. Tables, Formulæ, Graphics, and Notated Music* and *17. Linking, Segmentation, and Alignment*.

Next, section *3.11. Reference Systems*, describes methods of encoding within a text the conventional system or systems used when making references to the text. Some reference systems have attained canonical authority and should be recorded to make the text useable in normal work; in other cases, a convenient reference system should be created by the creator or analyst of an electronic text.

Like lists and notes, the bibliographic citations discussed in section *3.12. Bibliographic Citations and References*, may be regarded as structural elements in their own right. A range of possibilities is presented for the encoding of bibliographic citations or references, which may be treated as simple phrases within a running text, or as highly-structured components suitable for inclusion in a bibliographic database.

Additional elements for the encoding of passages of verse or drama (whether prose or verse) are discussed in section *3.13. Passages of Verse or Drama*.

The chapter concludes with a technical overview of the structure and organization of the module described here. This should be read in conjunction with chapter *1. The TEI Infrastructure*, describing the structure of the TEI document type definition.

## 3.1 Paragraphs

Paragraphs in modern printed or online text are typically visually offset with whitespace or an indented first line. But paragraphs are not simply blocks of text. The paragraph may be thought of as a mid-level unit of sense, a coherent grouping of sentences. Paragraphs may, in turn, be grouped into larger divisions, such as chapters. Because it is a unit of sense rather than simply a block of text, the <p> element in TEI may contain other structures displayed as blocks, such as lists or quotes. This distinguishes it from the p element in HTML, which is primarily a block of text, and from the <ab> (anonymous block) element described in *17.3. Blocks, Segments, and Anchors* which may be used as an alternative to the paragraph in cases that require a chunk-level container which is not necessarily a sense unit and which may have different structural properties.

The paragraph can be understood in the context of the distinct forms of textual division discussed in *1.3.2.1. Informal Element Classifications* : *chunk*, *phrase*, and *inter*-level. Chunk-level elements are paragraphs and other elements which have similar structural properties. Phrase-level elements must be entirely contained within a paragraph or other chunk-level structure. This type includes emphasized or quoted phrases, names, dates, etc. Inter-level elements can appear either within a paragraph or between paragraphs, and include bibliographic citations, notes, and lists. The <ab> (anonymous block) element is an alternative to <p> which is useful in cases where paragraphs are not present, but chunk-level organization is still needed. <ab> may contain other <ab>s, may use the *type* attribute, and does not necessarily represent a coherent set of statements.

Paragraphs can contain many of the other elements described within this chapter, as well as other elements which are specific to individual text types. Because paragraphs may appear in different customizations, their possible contents may vary in different kinds of documents. In particular, additional elements not listed in this chapter may appear in paragraphs. However, the elements described in this chapter are available in all kinds of text unless they are excluded by a customization.

The paragraph is marked using the <p> element:

`<p>` (paragraph) marks paragraphs in prose.

If a consistent internal subdivision of paragraphs is desired, the <s> or <seg> ('segment') elements may be used, as discussed in chapters *17. Linking, Segmentation, and Alignment* and *18. Simple Analytic Mechanisms* respectively. More usually, however, paragraphs have no firm internal structure, but contain prose encoded as a mix of characters, entity references, phrases marked as described in the rest of this chapter, and embedded elements like lists, figures, or tables.

The following example demonstrates a basic application of the <p> tag for an indented passage in a printed book:

```
<body>
  <p>It is in vain to say human beings ought to be satisfied with tranquillity:
     they must have action; and they will make it if they cannot find it. Millions are
     condemned to a stiller doom than mine, and millions are in silent revolt against
     their lot. Nobody knows how many rebellions besides political rebellions ferment
     in the masses of life which people earth. Women are supposed to be very calm
     generally: but women feel just as men feel; they need exercise for their
faculties,
     and a field for their efforts, as much as their brothers do; they suffer from
     too rigid a restraint, too absolute a stagnation, precisely as men would suffer;
and
     it is narrow-minded in their more privileged fellow-creatures to say that they
ought to
     confine themselves to making puddings and knitting stockings, to playing on the
piano
     and embroidering bags. It is thoughtless to condemn them, or laugh at them, if
they
     seek to do more or learn more than custom has pronounced necessary for their
sex.</p>
</body>
```

This news story shows typically short journalistic paragraphs:

```
<dateline>SARAJEVO, Bosnia and Herzegovina, April 19</dateline>
<p>Serbs seized more territory in this struggling new country today as
the United States Air Force ended a two-day airlift of humanitarian
aid into the capital, Sarajevo.</p>
<p>International relief workers called on European Community nations
to step up their humanitarian aid to the former Yugoslav republic,
in conjunction with new American aid flights if necessary.</p>
<p>A special envoy from the European Community, Colin Doyle, harshly
condemned the decision by Serbs to shell Sarajevo on Saturday night
during a visit to the Bosnian capital by a senior American official,
Deputy Assistant Secretary of State Ralph R. Johnson.</p>
<p>...</p>
```

The following extract from a Russian fairy tale demonstrates how other phrase level elements (in this case <q> elements representing direct speech; see section *3.3.3. Quotation*) may be nested within, but not across, paragraphs:

```
<p>A fly built a castle, a tall and mighty castle.
There came to the castle the Crawling Louse. <q>Who,
    who's in the castle? Who, who's in your house?</q>
said the Crawling Louse. <q>I, I, the Languishing Fly.
    And who art thou?</q>
  <q>I'm the Crawling Louse.</q>
</p>
<p>Then came to the castle the Leaping Flea. <q>Who,
    who's in the castle?</q> said the Leaping Flea. <q>I,
    I, the Languishing Fly, and I, the Crawling Louse. And
    who art thou?</q>
  <q>I'm the Leaping Flea.</q>
</p>
<p>Then came to the castle the Mischievous Mosquito.
<q>Who, who's in the castle?</q> said the Mischievous
Mosquito. <q>I, I, the Languishing Fly, and I, the
    Crawling Louse, and I, the Leaping Flea. And who art
    thou?</q>
  <q>I'm the Mischievous Mosquito.</q>
</p>
```

## 3.2 Treatment of Punctuation

Punctuation marks cause two distinct classes of problem for text markup: the marks may not be available in the character set used, and they may be significantly ambiguous. To some extent, the availability of the Unicode character set addresses the first of these problems, since it provides specific code points for most punctuation marks, and also the second to the extent that it distinguishes glyphs (such as stop, comma, and hyphen) which are used with different functions. Where punctuation itself is the subject of study, the element <pc> (punctuation character) may be used to mark it explicitly, as further discussed in *18.1.2. Below the Word Level*. Where the character used for a punctuation mark is not available in Unicode, the <g> element and other facilities described in chapter *5. Characters, Glyphs, and Writing Modes* may also be used to mark its presence.

### 3.2.1 Functions of Punctuation

Punctuation is itself a form of markup, historically introduced to provide the reader with an indication about how the text should be read. As such, it is unsurprising that encoders will often wish to encode directly the purpose for which punctuation was provided, as well as, or even instead of, the punctuation itself. We discuss some typical cases below.

The *Full stop (period)* may mark (orthographic) sentence boundaries, abbreviations, decimal points, or serve as a visual aid in printing numbers. These usages can be distinguished by tagging S-units, abbreviations, and numbers, as

described in sections *17.3. Blocks, Segments, and Anchors*, *3.6.5. Abbreviations and Their Expansions*, and *3.6.3. Numbers and Measures* respectively. However, there are independent reasons for tagging these, whether or not they are marked by full stops, and the polysemy of the full stop itself is perhaps no different from that of any other character in the writing system.

The *Question mark* and *exclamation mark* usually mark the end of orthographic sentences, but may also be used as a mid-sentence comment by the author (*!* to express surprise or some other strong feeling, *?* to query a word or expression or mark a sentence as dubious in linguistic discussion). Such usages may be distinguished by marking S-units, in which case the mid-sentence uses of these punctuation marks may be left unmarked, or tagged using the <pc> element discussed in *18.1. Linguistic Segment Categories*.

*Dashes* are used for a variety of purposes: as a mark of omission, insertion, or interruption; to show where a new speaker takes over (in dialogue); or to introduce a list item. In the latter two cases particularly, it is clearly desirable to mark the function as well as its rendition using the elements <q> or <item>, on which see section *3.3.3. Quotation*, and section *3.8. Lists*, respectively.

*Quotation marks* may be removed from text contained by <q> or <quote> elements on editorial grounds, or they may be marked in a variety of ways; see the discussion of quotation and related features in section *3.3.3. Quotation*.

*Apostrophes* should be distinguished from single quote marks. As with hyphens, this disambiguation is best performed by selecting the appropriate Unicode character, though it may also be represented by using appropriate XML markup for quotations as suggested above. However, apostrophes have a variety of uses. In English they mark contractions, genitive forms, and (occasionally) plural forms. Full disambiguation of these uses belongs to the level of linguistic analysis and interpretation.

*Parentheses* and other marks of suspension such as dashes or ellipses are often used to signal information about the syntactic structure of a text fragment. Full disambiguation of their uses also belongs to the level of linguistic analysis and interpretation, and will therefore need to use the mechanisms discussed in chapter *18. Simple Analytic Mechanisms*.

Where punctuation marks are disambiguated by tagging their assumed function in the text (for example, quotation), it may be debated whether they should be excluded or left as part of the text. In the case of quotation marks, it may be more convenient to distinguish opening from closing marks simply by using the appropriate Unicode character than to use the <q> element, with or without an indication of rendition.

Where segmentation of a text is performed automatically, the accuracy of the result may be considerably enhanced by a first pass in which the function of different punctuation characters is explicitly marked. This need not be done for all cases, but only where the structural function of the punctuation markup (for example as a word or phrase delimiter) is ambiguous. Thus, dots indicating abbreviation might be distinguished from dots indicating sentence end, and exclamation or question marks internal to a sentence distinguished from those which terminate one. Furthermore, when encoding historical materials, it may be considered essential to retain the original punctuation, whether by using an appropriate character code, if this is available (or using the <g> element where it is not) or by an explicit encoding using <pc>. The particular method adopted will vary depending upon the feature concerned and upon the purpose of the project.

### 3.2.2 Hyphenation

Hyphenation as a phenomenon is generally of most concern when producing formatted text for display in print or on screen: different languages and systems have developed quite sophisticated sets of rules about where hyphens may be introduced and for what reason. These generally do not concern the text encoder, since they belong to the domain of formatting and will generally be handled by the rendition software in use. In this section, we discuss issues arising from the appearance of hyphens in pre-existing formatted texts which are being re-encoded for analysis or other processing. Unicode distinguishes four characters visually similar to the hyphen, including the undifferentiated hyphen-minus (U+002D) which is retained for compatibility reasons. The hard hyphen (U+2010) is distinguished from the minus sign (U+2212) which is for use in mathematical expressions, and also from the soft hyphen (U+00AD) which may appear in 'born digital' documents to indicate places where it is acceptable to insert a hyphen when the document is formatted.

Historically, the hard hyphen has been used in printed or manuscript documents for two distinct purposes. In many languages, it is used between words to show that they function as a single syntactic or lexical unit. For example, in French, *est-ce que*; in English *body-snatcher*, *tea-party* etc. It may also have an important role in disambiguation (for example, by distinguishing say a *man-eating fish* from a *man eating fish*). Such usages, although possibly problematic when a linguistic analysis is undertaken, are not generally of concern to text encoders: the hyphen character is

usually retained in the text, because it may be regarded as part of the way a compound or other lexical item is spelled. Deciding whether a compound is to be decomposed into its constituent parts, and if so how, is a different question, involving consideration of many other phenomena in addition to the simple presence of a hyphen.

When it appears at the end of a printed or written line however, the hard hyphen generally indicates that—contrary to what might be expected—a word is not yet complete, but continues on the next line (or over the next page or column or other boundary). The hyphen character is not, in this case, part of the word, but just a signal that the word continues over the break. Unfortunately, few languages distinguish these two cases visually, which necessarily poses a problem for text encoders. Suppose, for example, that we wish to investigate a diachronic English corpus for occurrences of *tea-pot* and *teapot*, to find evidence for the point at which this compound becomes lexicalized. Any case where the word is hyphenated across a line division, like this:

```
  tea-
pot
```

is ambiguous: there is no simple way of deciding which of the two spellings was intended.
As elsewhere, therefore, encoders have a range of choices:

- They may decide simply to remove any end-of-line hyphenation from the encoded text, on the grounds that its presence is purely a secondary matter of formatting. This will obviously apply also if line endings are themselves regarded as unimportant.
- Alternatively, they may decide to record the presence of the hyphen, perhaps on the grounds that it provides useful morphological information; perhaps in order to retain information about the visual appearance of the original source. In either case, they need to decide whether to record it explicitly, by including an appropriate punctuation character in the text data, or implicitly by supplying an appropriate symbolic value for one or more of the attributes on the <lb> or other milestone element used to record the fact of the line division. If the hyphen is included in the character data of the TEI document, it might be marked up using the <pc> (punctuation character) tag, which allows the encoder to express information about its function as a separator, through the *force* attribute. For example, the example above could be encoded with a *force* value of "inter" to indicate that the punctuation mark may or may not be a word separator (See also *18.1.2. Below the Word Level*).

```
tea<pc force="inter">-</pc>pot
```

A similar range of possibilities applies equally to the representation of other common punctuation marks, notably quotation marks, as discussed in *3.3.3. Quotation*.

The 'text data' of which XML documents are composed is decomposable into smaller units, here called *orthographic tokens*, even if those units are not explicitly indicated by the XML markup. The ambiguity of the end-of-line hyphen also causes problems in the way a processor identifies such tokens in the absence of explicit markup. If token boundaries are not explicitly marked (for example using the <seg> or <w> elements), for most languages a processor will rely on character class information to determine where they are to be found: some punctuation characters are considered to be word-breaking, while others are not. In XML, the newline character in text data is a kind of whitespace, and is therefore word breaking. However, it is generally unsafe to assume that whitespace adjacent to markup tags will always be preserved, and it is decidedly unsafe to assume that markup tags themselves are equivalent to whitespace.

The <lb>, <pb>, and <cb> elements are notable exceptions to this general rule, since their function is precisely to represent (or replace) line, page, or column beginnings, which, as noted above, are generally considered to be equivalent to whitespace. These elements provide a more reliable way of preserving the lineation, pagination, etc. of a source document, since the encoder should not assume that (untagged) line beginnings etc. in an XML source file will necessarily be preserved.

To control the intended tokenization, the encoder may use the *break* attribute on such elements to indicate whether or not the element is to be regarded as equivalent to whitespace. This attribute can take the values yes or no to indicate whether or not the element corresponds with a token boundary. The value maybe is also available, for cases where the encoder does not wish (or is unable) to determine whether the orthographic token concerned is broken by the line ending.

As a final complication, it should be noted that in some languages, particularly German and Dutch, the spelling of a word may be altered in the presence of end of line hyphenation. For example, in Dutch, the word *opaatje* ( *granddad*), occurring at the end of a line may be hyphenated as *opa-tje*, with a single letter a. An encoder wishing to preserve the original form of this orthographic token in a printed text while at the same time facilitating its recognition as the word *opaatje* will therefore need to rely on a more sophisticated process than simply removing the hyphen. This is however essentially the same as any other form of normalization accompanying the recognition of variations in spelling or morphology: as such it may be encoded using the <choice> element discussed in *3.5. Simple Editorial Changes*, or the more sophisticated mechanisms for linguistic analysis discussed in chapter *18. Simple Analytic Mechanisms*.

## 3.3 Highlighting and Quotation

This section deals with a variety of textual features, all of which have in common that they are frequently realized in conventional printing practice by the use of such features as underlining, italic fonts, or quotation marks, collectively referred to here as *highlighting*. After an initial discussion of this phenomenon and alternate approaches to encoding it, this section describes ways of encoding the following textual features, all of which are conventionally rendered using some kind of highlighting:

- emphasis, foreign words and other linguistically distinct uses of highlighting
- representation of speech and thought, quotation, etc.
- technical terms, glosses, etc.

### 3.3.1 What Is Highlighting?

By *highlighting* we mean the use of any combination of typographic features (font, size, hue, etc.) in a printed or written text in order to distinguish some passage of a text from its surroundings.[44] The purpose of highlighting is generally to draw the reader's attention to some feature or characteristic of the passage highlighted; this section describes the elements recommended by these Guidelines for the encoding of such textual features.

In conventionally printed modern texts, highlighting is often employed to identify words or phrases which are regarded as being one or more of the following:

- distinct in some way—as foreign, dialectal, archaic, technical, etc.
- emphatic, and which would for example be stressed when spoken
- not part of the body of the text, for example cross-references, titles, headings, labels, etc.
- identified with a distinct narrative stream, for example an internal monologue or commentary.
- attributed by the narrator to some other agency, either within the text or outside it: for example, direct speech or quotation.
- set apart from the text in some other way: for example, proverbial phrases, words mentioned but not used, names of persons and places in older texts, editorial corrections or additions, etc.

The textual functions indicated by highlighting may not be rendered consistently in different parts of a text or in different texts. (For example, a foreign word may appear in italics if the surrounding text is in roman, but in roman if the surrounding text is in italics.) For this reason, these Guidelines distinguish between the encoding of rendering itself and the encoding of the underlying feature expressed by it.

Highlighting as such may be encoded by using one of the global attributes *rend*, *rendition*, or *style* (see further *1.3.1.1. Global Attributes*). This allows the encoder both to specify the function of a highlighted phrase or word, by selecting the appropriate element described here or elsewhere in these Guidelines, and to further describe the way in which it is highlighted, by means of an attribute. If the encoder wishes to offer no interpretation of the feature underlying the use of highlighting in the source text, then the <hi> element may be used, which indicates only that the text so tagged was highlighted in some way.

`<hi>` (highlighted) marks a word or phrase as graphically distinct from the surrounding text, for reasons concerning which no claim is made.

The <hi> element is provided by the `model.hiLike` class.

[44] Although the way in which a spoken text is performed, (for example, the voice quality, loudness, etc.) might be regarded as analogous to 'highlighting' in this sense, these Guidelines recommend distinct elements for the encoding of such 'highlighting' in spoken texts. See further section *8.3.6. Shifts*.

The possible values carried by the *rend* attribute are not formally defined in this version of the Guidelines. It may be used to document any peculiarity of the way a given segment of text was rendered in the original source text, and may thus express a very large range of typographic or other features, by no means restricted to typeface, type size, etc. The *style* attribute, by contrast, defines the way the source text was rendered using a formally defined style language, such as the W3C standard Cascading Stylesheet Language (**?**). The complementary *rendition* attribute is used to point to one or more fragments expressed using such a language which have been predefined in the TEI header using the <rendition> element discussed in section *2.3.4. The Tagging Declaration*.

Where it is both appropriate and feasible, these Guidelines recommend that the textual feature marked by the highlighting should be encoded, rather than just the simple fact of the highlighting. This is for the following reasons:

- the same kind of highlighting may be used for different purposes in different contexts
- the same textual function may be highlighted in different ways in different contexts
- for analytic purposes, it is in general more useful to know the intended function of a highlighted phrase than simply that it is distinct.

In many, if not most, cases the underlying function of a highlighted phrase will be obvious and non-controversial, since the distinctions indicated by a change of highlighting correspond with distinctions discussed elsewhere in these Guidelines. The elements available to record such distinctions are, for the most part, members of the `model.emphLike` class. This and the `model.hiLike` class mentioned above constitute the `model.highlighted` class, which is a phrase level class. Members of this class may appear anywhere within paragraph level elements.

The distinction between the two classes is simple, and typified by the two elements <hi> and <emph>: the former marks simply that a passage is typographically distinct in some way, while the latter asserts that a passage is linguistically emphasized for some purpose. These two properties, though often combined, are not identical. It should however be recognized, however, that cases do exist in which it is not economically feasible to mark the underlying function (e.g. in the preparation of large text corpora), as well as cases in which it is not intellectually appropriate (as in the transcription of some older materials, or in the preparation of material for the study of typographic practice). In such cases, the <hi> element or some other element from the `model.hiLike` class should be used.

Elements which are sometimes realized by typographic distinction but which are not discussed in this section include <title> (discussed in section *3.12. Bibliographic Citations and References*) and <name> (discussed in section *3.6.1. Referring Strings*).

## 3.3.2 Emphasis, Foreign Words, and Unusual Language

This subsection discusses the following elements:

`<foreign>` (foreign) identifies a word or phrase as belonging to some language other than that of the surrounding text.

`<emph>` (emphasized) marks words or phrases which are stressed or emphasized for linguistic or rhetorical effect.

`<distinct>` identifies any word or phrase which is regarded as linguistically distinct, for example as archaic, technical, dialectal, non-preferred, etc., or as forming part of a sublanguage.

These elements are all members of the `model.emphLike` class.

### 3.3.2.1 Foreign Words or Expressions

Words or phrases which are not in the main language of the text should be tagged as such, at least where the fact is indicated in the text. Where the word or phrase concerned is already distinguished from the rest of the text by virtue of its function (for example, because it is a name, a technical term, a quotation, a mentioned word, etc.) then the global *xml:lang* attribute should be used to specify additionally that its language distinguishes it from the surrounding text. Any element in the TEI scheme may take an *xml:lang* attribute, which specifies both the writing system and the language used by its content (see sections *1.3.1.1.2. Language Indicators* and *vi.1 Language Identification* for discussion of this attribute and its values respectively). Where there is no other applicable element, the element <foreign> may be used to provide a peg onto which the *xml:lang* may be attached.

```
<q>Aren't you confusing <foreign  xml:lang="la">post hoc</foreign> with
<foreign  xml:lang="la">propter
    hoc</foreign>?</q> said the Bee Master.
<q>Wax-moth only succeed when
weak bees let them in.</q>
```

The <foreign> element should not be used to represent foreign words which are mentioned or glossed within the text: for these use the appropriate element from section *3.4.1. Terms and Glosses* below. Compare the following example sentences:

```
John eats a <foreign  xml:lang="fr">croissant</foreign> every morning.
```

```
<mentioned  xml:lang="fr">Croissant</mentioned> is difficult to
pronounce with your mouth full.
```

```
A <term  xml:lang="fr">croissant</term> is a crescent-shaped
piece of light, buttery, pastry that is usually eaten for
breakfast, especially in France.
```

Elements which do not explicitly state the language of their content by means of an *xml:lang* attribute are understood to inherit a value for it from their parent element. In the general case, therefore, it is recommended practice to supply a default value for *xml:lang* on the root <TEI> or <text> element, as further discussed in section *1.3.1.1.2. Language Indicators*

#### 3.3.2.2 Emphatic Words and Phrases

The <emph> element is provided to mark words or phrases which are *linguistically* emphatic or stressed. Text which is only typographically 'emphasized' falls into the class of highlighted text, and may be tagged with the <hi> element. In printed works, emphasis is generally indicated by devices such as the use of an italic font, a large typeface, or extra wide letter spacing; in manuscripts and typescripts, it is usually indicated by the use of underlining. As the following examples demonstrate, an encoder may choose whether or not to make explicit the particular type of rendition associated with the emphasis. If a source text consistently renders a particular feature (e.g. emphasis or words in foreign languages) in a particular way, the rendering associated with that feature may be described in the TEI header using the <rendition> element. The *rend*, *rendition*, or *style* attributes may then be used to describe examples which deviate from the norm. For example, assuming that the TEI header has defined a default rendering for the <emph> element, the following encoding would use it:

```
<q>Sex, sir, is <emph>purely</emph> a question
of appetite!</q> Tarr exclaimed.
```

If on the other hand no such default has been defined for the element, the encoder may specify it informally using the *rend* attribute:

```
<q>What it all comes to is this,</q> he said.

<q>
  <emph  rend="italic">What does Christopher Robin do in the morning
    nowadays?</emph>
</q>
```

If the encoder wishes to express information about the rendition used in the source using a formal language such as CSS, then the *style* attribute can be used in a similar way:

```
<q>What it all comes to is this,</q> he said.

<q>
  <emph  style="font-style: italic">What does Christopher Robin do in
     the morning nowadays?</emph>
</q>
```

In cases where the rendition of a source needs to be indicated several times in a document, it may be more convenient to provide a default value using the <rendition> element in the header. If a small number of distinct values are required, it may also be convenient to define them all by means of a series of <rendition> elements which can then be referenced from the elements in question by means of the global *rendition* attribute:

```
<l>Here Thou, great <name  rendition="#italic">Anna</name>!
whom three Realms obey,</l>
<l>Doth sometimes Counsel take —
and sometimes <emph  rendition="#italic">Tea</emph>.</l>
<!-- in the header ... -->
<rendition  xml:id="italic"  scheme="css">font-style: italic</rendition>
```

Further information on the use of the <rendition> element is provided at *2.3.4. The Tagging Declaration*.

The <hi> element is used to mark words or phrases which are highlighted in some way, but for which identification of the intended distinction is difficult, controversial, or impossible. It enables an encoder simply to record the fact of highlighting, possibly describing it by the use of a *rend*, *style*, or *rendition* attribute, as discussed above, without however taking a position as to the function of the highlighting. This may also be useful if the text is to be processed in two stages: representing simply typographic distinctions during a first pass, and then replacing the <hi> elements with more specific elements in a second pass.

Some simple examples:

```
<hi  rend="gothic">And this Indenture further witnesseth</hi>
that the said <hi  rend="italic">Walter Shandy</hi>, merchant,
in consideration of the said intended marriage ...
```

In this example, the first highlighted phrase uses black letter or gothic print to mimic the appearance of a legal document, and italic to mark *Walter Shandy* as a name. In a second pass, the elements <head> or <label> might be appropriate for the first use, and the element <name> for the second.

```
The heaviest rain, and snow, and hail, and
sleet, could boast of the advantage over him in only one respect. They
often <hi  rend="quoted">came down</hi> handsomely, and Scrooge never
did.
```

In this example, the phrase *came down* uses inverted commas to indicate a play on words.[45] In a second pass, the element <soCalled> might be preferred as a means of indicating that the narrator is distancing himself from this usage.

[45] The Oxford English Dictionary documents the phrase *to come down* in the sense 'to bring or put down; *esp.* to lay down money; to make a disbursement' as being in use, mostly in colloquial or humorous contexts, from at least 1700 to the latter half of the 19th century.

#### 3.3.2.3 Other Linguistically Distinct Material

For some kinds of analysis, it may be desirable to encode the linguistic distinctiveness of words and phrases with more delicacy than is allowed by the <foreign> element. The <distinct> element is provided for this purpose. Its attributes allow for additional information characterizing the nature of the linguistic distinction to be made in two distinct ways: the *type* attribute simply assigns a user-defined code of some kind to the word or phrase which assigns it to some register, sub-language, etc. No recommendations as to the set of values for this attribute are provided at this time, as little consensus exists in the field.

Alternatively, the remaining three attributes may be used in combination to place a word or phrase on a three-dimensional scale sometimes used in descriptive linguistics, as for example in Mattheier et al, 1988. The *time* attribute places a word or phrase *diachronically*, for example as archaic, old-fashioned, contemporary, futuristic, etc.; the *space* attribute places a word or phrase *diatopically*, that is, with respect to a geographical classification, for example as national, regional, international, etc.; the *social* attribute places a word or phrase *diastratically*, that is, with respect to a social classification, for example as technical, polite, impolite, restricted, etc. Again, no recommendations are made for the values of these attributes at this time; the encoder should provide a description of the scheme used in the appropriate section of the header (see section *2.3. The Encoding Description*).

Examples:

```
Next morning a boy in that dormitory confided to his
bosom friend, a <distinct  type="psSlang">fag</distinct> of
Macrea's, that there was trouble in their midst which
King <distinct  type="archaic">would fain</distinct> keep
secret.
```

```
Next morning a boy in that dormitory confided to his
bosom friend, a
<distinct  time="1900"  space="GB"
  social="publicschool">fag</distinct>
of Macrea's, that there was trouble in their midst which
King <distinct  time="archaic">would fain</distinct> keep
secret.
```

Where more complex (or more rigorous) interpretive analyses of the associations of a word are required, the more detailed and general mechanisms described in chapter *19. Feature Structures* should be preferred to these simple characterizations. It may also be preferable to record the kinds of analysis suggested here by means of the simple annotation element <note> described in section *3.9. Notes, Annotation, and Indexing*, or the  element described in section *18.3. Spans and Interpretations*.

### 3.3.3 Quotation

One form of presentational variation found particularly frequently in written and printed texts is the use of quotation marks. As with the typographic variations discussed in the preceding section, it is generally helpful to separate the encoding of the underlying textual feature (for example, a quotation or a piece of direct speech) from the encoding of its rendering (for example, the use of a particular style of quotation marks).

This section discusses the following elements, all of which are often rendered by the use of quotation marks:

`<q>` (quoted) contains material which is distinguished from the surrounding text using quotation marks or a similar method, for any one of a variety of reasons including, but not limited to: direct speech or thought, technical terms or jargon, authorial distance, quotations from elsewhere, and passages that are mentioned but not used.

`<said>` (speech or thought) indicates passages thought or spoken aloud, whether explicitly indicated in the source or not, whether directly or indirectly reported, whether by real people or fictional characters.

`@direct` may be used to indicate whether the quoted matter is regarded as direct or indirect speech.

`@aloud` may be used to indicate whether the quoted matter is regarded as having been vocalized or signed.

`<quote>` (quotation) contains a phrase or passage attributed by the narrator or author to some agency external to the text.

`att.global.source` provides attributes used by elements to point to an external source.

`@source` specifies the source from which some aspect of this element is drawn.

`<cit>` (cited quotation) contains a quotation from some other document, together with a bibliographic reference to its source. In a dictionary it may contain an example text with at least one occurrence of the word form, used in the sense being described, or a translation of the headword, or an example.

`<mentioned>` marks words or phrases mentioned, not used.

`<soCalled>` (so called) contains a word or phrase for which the author or narrator indicates a disclaiming of responsibility, for example by the use of scare quotes or italics.

The elements <mentioned> and <soCalled> are members of the class `model.emphLike` while <q> stems from `model.hiLike`; the element <said> is a member of the class `model.attributable` in its own right, while <cit> and <quote> are members of `model.quoteLike`, a subclass of `model.attributable`. This class is a subclass of `model.inter`; hence all of these elements are permitted both within and between paragraph-level elements.

The most common and important use of quotation marks is, of course, to mark *quotation*, by which we mean simply any part of the text which the author or narrator wishes to attribute to some agency other than the narrative voice. The <q> element may be used if no further distinction beyond this is judged necessary. If it is felt necessary to distinguish such passages further, for example to indicate whether they are regarded as speech, writing, or thought, either the *type* attribute or one of the more specialized elements discussed in this section may be used. For example, the element <quote> may be used for written passages cited from other works, or the element <said> for words or phrases represented as being spoken or thought by people or characters within the current work. The <soCalled> element is used for cases where the author or narrator distances him or herself from the words in question without however attributing them to any other voice in particular. The <mentioned> element is appropriate for a case where a word or phrase is being discussed in the body of a text rather than forming part of the text directly.

As noted above, if the distinction among these various reasons why a passage is offset from surrounding text cannot be made reliably, or is not of interest, then any representation of speech, thought, or writing may simply be marked using the <q> element.

Quotation may be indicated in a printed source by changes in type face, by special punctuation marks (single or double or angled quotes, dashes, etc.) and by layout (indented paragraphs, etc.), or it may not be explicitly represented at all. If these characteristics are of interest, one or other of the global *rend* or *rendition* attributes discussed in section *1.3.1.1. Global Attributes* may be used to record them.

Quotation marks themselves may, like other punctuation marks, be felt for some purposes to be worth retaining within a text, quite independently of their description by the *rend* attribute. This should generally be done using the appropriate Unicode character, or, if this is not possible, a numeric character reference (see *v.7.1 Character References*). If the encoder decides both to retain the quotation marks and to represent their function by means of an explicit tag such as <quote>, the quotation marks should be included within the element, rather than outside it, as in the first example below:

```
Adolphe se tourna vers lui :
<said>— Alors, Albert, quoi de neuf?</said>
<said>— Pas grand-chose.</said>
<said>— Il fait beau,</said> dit Robert.
```

Alternatively, since this use of the leading mdash is very common typographic practice, it may be considered unnecessary to retain it in the encoding. Its presence in the source might instead be signalled using one of the attributes *rend*, *style*, or *rendition*. This kind of rendering might be predefined using a <rendition> element, which can then be referenced using the *rendition* attribute as follows:

```
Adolphe se tourna vers lui :
<said rendition="#dashBefore">Alors,
Albert, quoi de neuf ?</said>
<said rendition="#dashBefore">Pas grand-chose.</said>
<said rendition="#dashBefore">Il fait beau,</said>
dit Robert.

<!-- ... within the header -->
<rendition xml:id="dashBefore"
  scope="before">content: '— '</rendition>
<!-- ... -->
<quotation marks="none"/>
```

Whatever policy is adopted, the encoder should document the decision in some way, for example by using the <quotation> element provided in the TEI header (see *2.3.3. The Editorial Practices Declaration*) to indicate that quotation marks have not been retained in the encoding; their presence in the source is implied by the *rendition* attribute values supplied.

Whether or not the quotation marks are suppressed, their presence and nature may be described using some appropriate set of conventions in the *rend* attribute. These conventions may be entirely idiosyncratic, and hence not necessarily useful for interchange, as in the following example:

```
<said rend="pre(‘) post(’)">Who-e debel
you?</said> — he at last said —
<said rend="pre(‘) post(’)">you no speak-e,
damme, I kill-e.</said> And so saying,
the lighted tomahawk began flourishing
about me in the dark.
```

Such passages might more effectively be encoded without loss of rendering information by using the *rendition* attribute and its associated <rendition> element as described in section *2.3.4.1. Rendition.* If the rendition of passages tagged as <said> is generally uniform throughout a text, then the <rendition> element may used to specify a default rendering, in which case the same section might simply be tagged:

```
<said>Who-e debel
you?</said> — he at last said —
<said>you no speak-e,
damme, I kill-e.</said> And so saying,
the lighted tomahawk began flourishing
about me in the dark.

<!-- in the header: -->
<tagsDecl partial="true">
  <rendition scheme="css" selector="said"
    scope="before">content:"‘";</rendition>
  <rendition scheme="css" selector="said"
    scope="after">content:"’";</rendition>
</tagsDecl>
```

As members of the `att.ascribed.directed` class, elements <said> and <q> share the following attributes:

`att.ascribed` provides attributes for elements representing speech or action that can be ascribed to a specific individual.

`@who` indicates the person, or group of people, to whom the element content is ascribed.

`att.ascribed.directed` provides attributes for elements representing speech or action that can be directed at a group or individual.

`@toWhom` indicates the person, or group of people, to whom a speech act or action is directed.

These may be used to make explicit who is speaking and who is being spoken to:

```
Adolphe se tourna vers lui :
<said  who="#Ado"  toWhom="#Alb">— Alors, Albert, quoi de neuf?</said>
<said  who="#Alb"  toWhom="#Ado">— Pas grand-chose.</said>
<said  who="#Rob">— Il fait beau,</said> dit Robert.

<!-- ... elsewhere in the document -->
<standOff>
  <listPerson  type="speakers">
    <person  xml:id="Ado">
      <persName>Adolphe</persName>
    </person>
    <person  xml:id="Alb">
      <persName>Albert</persName>
    </person>
    <person  xml:id="Rob">
      <persName>Robert</persName>
    </person>
  </listPerson>
</standOff>
```

The *who* attribute may be supplied whether or not an indication of the speaker is given explicitly in the text. Likewise, the *toWhom* attribute may be supplied to indicate the encoder's interpretation of the intended recipient of the speech. They may take the form (as above) of an abbreviated or normalized form of the speaker's name, but the role of these attributes is to act as a pointer to a location elsewhere in the text, or another document, where data about each party may be supplied. While these attributes could point to any source of information about the parties involved, the most appropriate place for such information is within the *participant description* component of the TEI header, as further discussed in *16.2.2. The Participant Description* but for simple cases like the above, a simple list of speakers located in the front or back matter of the text may suffice.

It may also be useful to distinguish representations of speech from representations of thought, in modern printed texts often indicated by a change of typeface. The *aloud* attribute is provided for this purpose, as in this example:

```
<said  aloud="true">Oh yes,</said> said Henry,
<said  aloud="false">I mean
Gordon Macrae, for example…</said>
<said  aloud="false">Jungian
Analyst with Winebox! That's what you called him, you callous bastard,
didn't you? Eh? Eh?</said>
```

Quoted matter may be embedded within quoted matter, as when one speaker reports the speech of another:

```
<said  who="#Wilson">Spaulding, he came down into the office just this day
eight weeks with this very paper in his hand, and he says:—
<said  who="#WilsonSpaulding">I wish to the Lord, Mr. Wilson, that I was a
    red-headed man.</said>
</said>
<!-- ... -->
<list  type="speakers">
  <item  xml:id="Wilson">Wilson</item>
```

```
  <item  xml:id="WilsonSpaulding">Spaulding reported by Wilson</item>
<!-- ...-->
</list>
```

Direct speech nested in this way is treated in the same way as elsewhere: a change of rendition may occur, but the same element should be used. An encoder may however choose to distinguish between direct speech which contains quotations from extra-textual matter and direct speech itself, as in the following example:

```
<p>
  <said>The Lord! The Lord! It is Sakya Muni himself,</said> the lama half
sobbed; and under his breath began the wonderful Buddhist
invocation:-<said>
    <quote>
      <l>To Him the Way — the Law — Apart —</l>
      <l>Whom Maya held beneath her heart</l>
      <l>Ananda's Lord — the Bodhisat</l>
    </quote>
     And He is here! The Most Excellent Law is here also. My
     pilgrimage is well begun. And what work! What work!</said>
</p>
```

Quotations from other works are often accompanied by a reference to their source. The <cit> element may be used to group together the quotation and its associated bibliographic reference, which should be encoded using the elements for bibliographic references discussed in section *3.12. Bibliographic Citations and References*, as in the following example.

```
<div  xml:id="mm01"  type="chapter">
  <head>Chapter 1</head>
  <epigraph>
    <cit>
      <quote>
        <l>Since I can do no good because a woman</l>
        <l>Reach constantly at something that is near it.</l>
      </quote>
      <bibl>
        <title>The Maid's Tragedy</title>
        <author>Beaumont and Fletcher</author>
      </bibl>
    </cit>
  </epigraph>
  <p>Miss Brooke had that kind of beauty which seems to be thrown into
     relief by poor dress...</p>
</div>
```

Like other bibliographic references, the citation associated with a quotation may be represented simply by a cross-reference, as in this example:

```
Lexicography has shown little sign of being affected by the
work of followers of J.R. Firth, probably best summarized
in his slogan, <cit>
  <quote>You shall know a word by the company it keeps.</quote>
  <ref>(Firth, 1957)</ref>
</cit>
```

It is also common for quotations to be separated from their bibliographic reference by intervening text, which makes the use of <cit> impractical. In such circumstances, the quotation can be linked to a bibliographical reference using *source*:

```
<bibl  xml:id="tlk_36">Tolkien (1936)</bibl> tells us that
<quote  source="#tlk_36">
  <title>Beowulf</title> is in fact so interesting as
poetry, in places poetry so powerful, that this quite
overshadows the historical content
</quote>.
```

*source* could also be used to point to a complete external bibliographic reference in a <listBibl> elsewhere in the document, or external to it.

Unlike most of the other elements discussed in this chapter, direct speech and quotations may frequently contain other high-level elements such as paragraphs or verse lines, as well as being themselves contained by such elements. Three possible solutions exist for this well-known structural problem:

- the quotation is broken into segments, each of which is entirely contained within a paragraph
- the quotation is marked up using stand-off markup
- the quotation boundaries are represented by empty segment boundary delimiter elements

For further discussion and several examples, see chapter *21. Non-hierarchical Structures*.

Finally, in this section, the element <soCalled> is provided for all cases in which quotation marks are used to distance the quoted text from the narrator or speaker. Common examples include the 'scare' quotes often found in newspaper headlines and advertising copy, where the effect is to cast doubts on the veracity of an assertion:

```
<head>PM dodges <soCalled>election threat</soCalled> in interview</head>
```

The same element should be used to mark a variety of special ironic usages. Some further examples follow:

```
He hated <soCalled>good</soCalled> books.
```

```
<soCalled>Croissants</soCalled> indeed! toast not good enough for you?
```

```
Although Chomsky's decision that all NL
sentences are finite objects was never justified by arguments from
the attested properties of NLs, it did have a certain
<soCalled>social</soCalled> justification. It was commonly assumed in
works on logic until fairly recently that the notion
<mentioned>language</mentioned> is necessarily restricted to finite
strings.
```

## 3.4 Terms and Glosses, Ruby Annotations, and Equivalents and Descriptions

This section describes a set of textual elements which are used to provide a gloss, alternate identification, or description of something.

### 3.4.1 Terms and Glosses

Technical terms are often italicized or emboldened upon first mention in printed texts; an explanation or gloss is sometimes given in quotation marks. Linguistic analyses conventionally cite words in languages under discussion in italics, providing a gloss immediately following marked with single quotation marks. Other texts in which individual words or phrases are *mentioned* (for example, as examples) rather than *used* may mark them either with italics or with quotation marks, and will gloss them less regularly.

`<term>` (term) contains a single-word, multi-word, or symbolic designation which is regarded as a technical term.

`<gloss>` (gloss) identifies a phrase or word used to provide a gloss or definition for some other word or phrase.

These elements are also members of the class `model.emphLike`.

A <term> may appear with or without a gloss, as may a <mentioned> element. Where the <gloss> is present, it may be linked to the term it is glossing by means of its *target* attribute. To establish such a link, the encoder should give an *xml:id* value to the <term> or <mentioned> element and provide that value, prefixed by a sharp sign, as the value of the *target* attribute on the <gloss> element. For example:

```
We may define <term xml:id="TDPv" rend="sc">discoursal point of view</term>
as
<gloss target="#TDPv">the relationship, expressed through discourse
structure, between the implied author or some other addresser,
and the fiction.</gloss>
```

```
<gloss rend="unmarked" target="#PRSR">A computational device that infers
structure from grammatical strings of words</gloss> is known as a
<term xml:id="PRSR">parser</term>, and much of the history of NLP over the
last 20 years has been occupied with the design of parsers.
```

In the absence of a <gloss> in the source text, a <term> can also be associated with a gloss or definition by means of its *ref* attribute, as in this imaginary example:

```
We discuss Leech's concept of <term ref="#TDPV2" rend="sc">discoursal point of
view</term> below.
```

Here, the value `#TDPV2` references some other XML element in the same document. This might be, for example, an entry in a list of technical vocabulary given in the document header, or in an appendix. This attribute could also reference an entry in some centrally maintained terminological database, perhaps using one of the pointing mechanisms discussed in section *17.2. Pointing Mechanisms*.

Note that the element <term> is intended for use with words or phrases identified as terminological in nature; where words or phrases are simply being cited, discussed, or glossed in a text, it will often be more appropriate to use the <mentioned> element, as in the following example:

```
There is thus a striking accentual difference between a verbal
form like <mentioned xml:id="cw234" xml:lang="grc">eluthemen</mentioned>
<gloss target="#cw234">we were released,</gloss> accented on the
second syllable of the word, and its participial derivative

<mentioned xml:id="cw235" xml:lang="grc">lutheis</mentioned>
<gloss target="#cw235">released,</gloss> accented on the last.
```

For technical terminology in particular, and generally in terminological studies, it may be useful to associate an instance of a term within a text with a canonical definition for it, which is stored either elsewhere in the same text

(for example in a glossary of terms) or externally, for example in a database, authority file, or published standard. The attributes *key* and *ref* discussed in section *3.6.1. Referring Strings* below are available on the <term> element for this purpose.

#### 3.4.1.1 Some Further Examples

As a simple example of the elements discussed here and in *3.3. Highlighting and Quotation*, consider the following sentence:

> On the one hand the *Nibelungenlied* is associated with the new rise of romance of twelfth-century France, the *romans d'antiquité*, the romances of Chrétien de Troyes, and the German adaptations of these works by Heinrich van Veldeke, Hartmann von Aue, and Wolfram von Eschenbach.

A first approximation to the encoding of this sentence might be simply to record the fact that the phrases printed above in italics are highlighted, as follows:

```
On the one hand the <hi  rend="italic">Nibelungenlied</hi> is
associated with the new rise of romance of twelfth-century France,
the <hi  xml:lang="fr"  rend="italic">romans d'antiquité</hi>,
the romances of Chrétien de Troyes, ...
```

This encoding would, however, lose the important distinction between an italicized title and an italicized foreign phrase. Many other phrases might also be italicized in the text, and a retrieval program seeking to identify foreign terms (for example) would not be able to produce reliable results by simply looking for italicized words. Where economic and intellectual constraints permit, therefore, it would be preferable to encode both the function of the highlighted phrases and their appearance, as follows:

```
On the one hand the <title  rend="italic">Nibelungenlied</title>
is associated with the new rise of romance of twelfth-century France,
the <foreign  rend="italic">romans d'antiquité</foreign>, the
romances of Chrétien de Troyes, ...
```

In this example, the decision as to which textual features are distinguished by the highlighting is relatively uncontroversial. As a less straightforward example, consider the use of italic font in the following passage:

> A pretty common case, I believe; in all *vehement* debatings. She says I am *too witty*; Anglicé, *too pert*; I, that she is *too wise*; that is to say, being likewise put into English, *not so young as she has been*: in short, she is grown so much into a *mother*, that she had forgotten she ever was a *daughter*. …

Clearly, the word *vehement* is not italicized for the same reason as the phrase *not so young as she has been*; the former is emphasized, while the latter is proverbial. It also provides an ironic gloss for the words *too wise*, in the same way as *too pert* glosses *too witty*. The glossed phrases are not, however, technical terms or cited words, but quoted phrases, as if the writer were putting words into her own and her mother's mouths. Finally, the words *mother* and *daughter* are apparently italicized simply to oppose them in the sentence; certainly they do not fit into any of the categories so far proposed as reasons for italicizing. Note also that the word *Anglicé* is not italicized although it is not generally considered an English word.

The following sample encoding for the above passage attempts to take into account all the above points:

```
A pretty common case, I believe; in all <emph>vehement</emph>
debatings. She says I am <q  rend="italic">too witty</q>;
<foreign  xml:lang="la"  rend="roman">Anglicé</foreign>,
<gloss  rend="italic">too pert</gloss>; I, that she is
<q  rend="italic"> too wise</q>; that is to say, being likewise
put into English, <gloss  rend="italic">not so young as she has
been</gloss>: in short, she is grown so much into a
<hi  rend="italic">mother</hi>, that she had forgotten she ever
was a <hi  rend="italic">daughter</hi>.
```

### 3.4.2 Ruby Annotations

The word *ruby* (or *rubi*) refers to a particular method of glossing runs of text which is common in East Asian scripts. In horizontally-oriented text, ruby annotations typically appear above the text being glossed, while in vertical runs of text they may appear to the left or right, or both, also oriented vertically. An English example of a ruby annotation might look like this:

In Japanese, furigana (振り仮名) ruby annotations are often used to provide pronunciation guidance for readers; characters from the largely phonetic hiragana or katakana syllabaries accompany Chinese characters, like this:

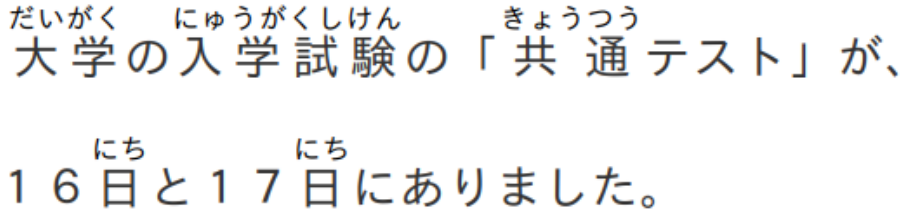



Figure 2: The first line of a news story from NHK News Web Easy intended for Japanese learners, in which every word composed of Chinese characters has a ruby gloss.

Pinyin ruby annotations are also used in Chinese to provide pronunciation guidance, and Zhuyin (注音) phonetic symbols (commonly known as *bopomofo*) are used in Taiwan for the same purpose.

The TEI schema provides many different ways of encoding glosses and annotations, from the simple and flexible <note> element to a native implementation of the Web Annotation Data Model (*17.11. Annotations*). However, ruby is a particular, distinct, and widely-used form of annotation that appears in script, print, calligraphy, and web pages, and the TEI therefore provides specific elements for it:

`<ruby>` (ruby container) contains a passage of base text along with its associated ruby gloss(es).

`<rb>` (ruby base) contains the base text annotated by a ruby gloss.

`<rt>` (ruby text) contains a ruby text, an annotation closely associated with a passage of the main text.

The <rt> element is a member of `att.placement`, and thus the *place* attribute may be used to indicate where the ruby gloss is with respect to the base text:

`att.placement` provides attributes for describing where on the source page or object a textual element appears.

`@place` specifies where this item is placed.

The most relevant suggested values of *place* for ruby text are above, below, left, and right.

In its simplest representation, a glossed form consists of an <rb> (ruby base) element containing the base form, an <rt> (ruby text) element containing the gloss, and a <ruby> element which wraps them together:

```
<p xml:lang="ja">
<!--…-->
  <ruby>
    <rb>大学</rb>
    <rt place="above">だいがく</rt>
  </ruby>
<!--…-->
</p>
```

In the example above, the word 大学 (*daigaku* = university) is provided with a phonation gloss in hiragana. The full gloss is applied to the complete word. However, it might instead be broken down by character:

```
<p  xml:lang="ja">
<!--…-->
  <ruby>
    <rb>大</rb>
    <rt  place="above">だい</rt>
  </ruby>
  <ruby>
    <rb>学</rb>
    <rt  place="above">がく</rt>
  </ruby>
<!--…-->
</p>
```

Here is a similar example from Taiwan using bopomofo (pinyin *píngzi* = bottle)[46]:

```
<p  xml:lang="zh-TW">
<!--…-->
  <ruby>
    <rb>瓶</rb>
    <rt  place="right">ㄆㄧㄥˊ</rt>
  </ruby>
  <ruby>
    <rb>子</rb>
    <rt  place="right">·ㄗ</rt>
  </ruby>
<!--…-->
</p>
```

Where *place* is not provided, the default assumption is that the ruby gloss is *above* where the text is horizontal, and to the *right* of the text where it is vertical. See *5.6. Writing Modes* for a detailed guide to writing modes and text directionality.

The same ruby base may be accompanied by more than one gloss. Here, the Japanese word 打球場 (dakyūba, or *billiard hall*) is glossed with two different pronunciations: 'biriādo' (its English equivalent) and 'dakyū', a phonation guide for the first two characters. This example is intriguing in that the right-side ruby glosses apply to the first and second characters respectively, but the left-side gloss applies to the whole word as a unit. We use this instance to exemplify multiple approaches to encoding the same phenomena, which may be appropriate for different projects or editorial preferences. First, using the same segmentation approach as demonstrated for 大学 above, but with nesting:

```
<p  style="writing-mode: vertical-rl"
  xml:lang="ja">
<!--…-->
  <ruby>
    <rb>
      <ruby>
        <rb>打</rb>
        <rt  place="right">ダ</rt>
      </ruby>
      <ruby>
        <rb>球</rb>
```

[46] Taken from Wikipedia.

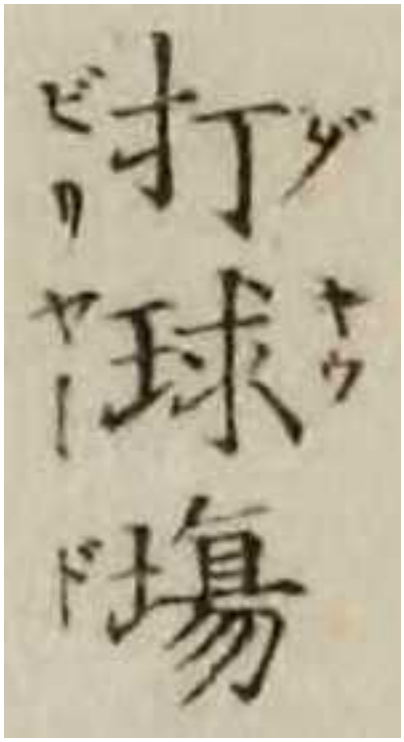


Figure 3: *Billiard hall* with two ruby glosses. 国文学研究資料館所蔵：：英国／龍動新繁昌記.

```
      <rt  place="right">キウ</rt>
     </ruby>
       場
   </rb>
   <rt  place="left">ビリヤード</rt>
  </ruby>
<!--…-->
</p>
```

We could also use a standoff approach with <anchor> elements and pointers:

```
<p  style="writing-mode: vertical-rl"
  xml:lang="ja">
<!--…-->
  <ruby>
    <rb>
      <anchor  xml:id="da"/>打
    <anchor  xml:id="kyuu"/>球
    <anchor  xml:id="ba"/>場
    <anchor  xml:id="owari"/>
    </rb>
    <rt  place="left"  from="#da"  to="#owari">ビリヤード</rt>
    <rt  place="right"  from="#da"  to="#kyuu">ダ</rt>
    <rt  place="right"  from="#kyuu"  to="#ba">キウ</rt>
  </ruby>
<!--…-->
</p>
```

Alternatively, if the encoding itself already includes segmentation below the word level, we can use the existing elements instead of adding <anchor>s:

```
<p  style="writing-mode: vertical-rl"
  xml:lang="ja">
<!--…-->
  <ruby>
    <rb  xml:id="dakyuuba">
      <c  xml:id="chr1">打</c>
      <c  xml:id="chr2">球</c>
      <c>場</c>
    </rb>
    <rt  place="left"  target="#dakyuuba">ビリヤード</rt>
    <rt  place="right"  target="#chr1">ダ</rt>
    <rt  place="right"  target="#chr2">キウ</rt>
  </ruby>
<!--…-->
</p>
```

Multiple ruby annotations can appear in the same alignment relative to the main text. For example, there might be two annotations on the right. We use this instance to exemplify two ways of encoding this phenomenon, dependent on the interpretation of the text. We could use a parallel approach where a sequence of two <rt> elements, both with the same value for the *place* attribute, are assumed to appear in their document sequence. The first of the two <rt> elements would appear to the right of the ruby base (<rb>), while the second would appear to the right of the first <rt> element [47]:

```
<p  style="writing-mode: vertical-rl"
  xml:lang="ja">
<!--…-->
  <ruby>
    <rb>ㄅ</rb>
    <rt  place="right">ㄅ</rt>
    <rt  place="right">博</rt>
  </ruby>
<!--…-->
</p>
```

Alternatively, a nested approach can be used. With nested <ruby> elements, the structure suggests that the second <rt> element is glossing both the <rb> and the <rt> element:

```
<p  style="writing-mode: vertical-rl"
  xml:lang="ja">
<!--…-->
  <ruby>
    <rb>
      <ruby>
        <rb>ㄅ</rb>
        <rt  place="right">ㄅ</rt>
      </ruby>
    </rb>
    <rt  place="right">博</rt>
  </ruby>
```

[47] Taken from Wikipedia.

```
<!--…-->
</p>
```

The <rt> element is a member of `att.written`:

`att.written` provides attributes to indicate the hand in which the content of an element was written in the source being transcribed.

*@hand* points to a <handNote> element describing the hand considered responsible for the content of the element concerned.

so the *hand* attribute can be used where the ruby text constitutes an annotation in a different hand from the ruby base:

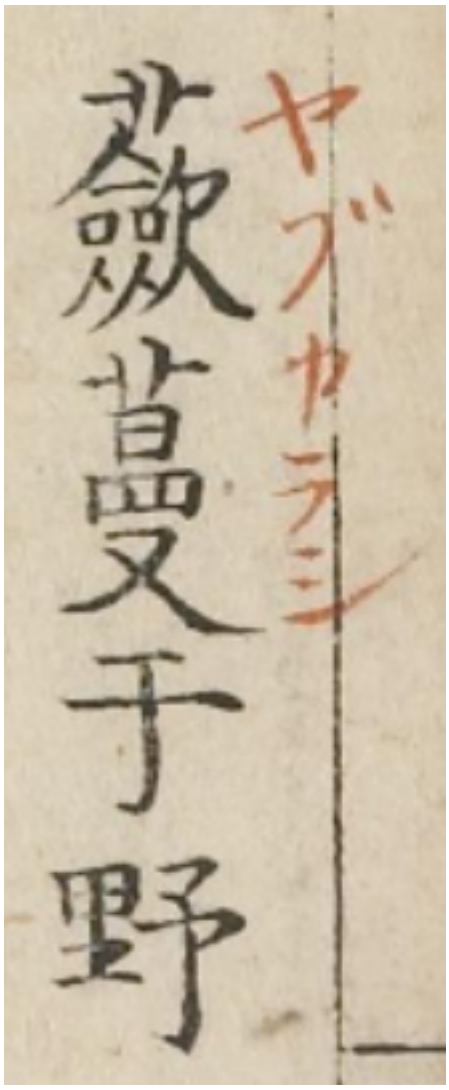


Figure 4: Base text with a ruby gloss in a different hand. リクシ ソウモク チョウジュウ チュウギョ ソ ズカイ。東京大学総合図書館 (General Library in the University of Tokyo, JAPAN).

```
<p  style="writing-mode: vertical-rl"
  xml:lang="ja">
```

```
<!--…-->
  <ruby>
    <rb>薮蔓</rb>
    <rt  hand="#MoriOhgai"  style="color: red;">ヤブカラシ</rt>
  </ruby>
于野

<!--…-->
</p>
```

The current support for ruby is rudimentary, and in future releases of the Guidelines we expect to see more development of these features and recommendations. While ruby is included for use with East Asian languages, encoders may find other contexts in which these elements are useful.

### 3.4.3 Equivalents and Descriptions

Another group of elements is used to supply different kinds of names for objects described by the TEI. Examples of this are documentation of elements, attributes, classes (and also attribute values where appropriate), and description of glyphs.

`<altIdent>` (alternate identifier) supplies the recommended XML name for an element, class, attribute, etc. in some language.

`<desc>` (description) contains a short description of the purpose, function, or use of its parent element, or when the parent is a documentation element, describes or defines the object being documented.

`<equiv>` (equivalent) specifies a component which is considered equivalent to the parent element, either by co-reference, or by external link.

- `@uri` (uniform resource identifier) references the underlying concept of which the parent is a representation by means of some external identifier.
- `@filter` references an external script which contains a method to transform instances of this element to canonical TEI.
- `@name` a single word which follows the rules defining a legal XML name (see `https://www.w3.org/TR/REC-xml/#dt-name`), naming the underlying concept of which the parent is a representation.
- `@predicate [att.predicate]` the condition under which the element bearing this attribute applies, given as an XPath predicate expression.

Along with the <gloss> element mentioned above, these elements constitute the `model.identSynonyms` class. They are described in more detail in *23.4.1. Description of Components*.

## 3.5 Simple Editorial Changes

As in editing a printed text, so in encoding a text in electronic form, it may be necessary to accommodate editorial comment on the text and to render account of any changes made to the text in preparing it. The tags described in this section may be used to record such editorial interventions, whether made by the encoder, by the editor of a printed edition used as a copy text, by earlier editors, or by the copyists of manuscripts.

The tags described here handle most common types of editorial intervention and stereotyped comment; where less structured commentary of other types is to be included, it may be marked using the <note> element described in section *3.9. Notes, Annotation, and Indexing*. Systematic interpretive annotation is also possible using the various methods described in chapter *17. Linking, Segmentation, and Alignment*. The examples given here illustrate only simple cases of editorial intervention; in particular, they permit economical encoding of a simple set of alternative readings of a short span of text. To encode multiple views of large or heterogeneous spans of text, the mechanisms described in chapter *17. Linking, Segmentation, and Alignment* should be used. To encode multiple witnesses of a particular text, a similar mechanism designed specifically for critical editions is described in chapter *13. Critical Apparatus*.

For most of the elements discussed here, some encoders may wish to indicate both a *responsibility*, that is, a code indicating the person or agency responsible for making the editorial intervention in question, and also

an indication of the degree of *certainty* which the encoder wishes to associate with the intervention. These requirements are served by the `att.global.responsibility` class, along with `att.global.source` and `att.dimensions`. Any of the elements discussed here thus may potentially carry any of the following optional attributes:

`att.global.responsibility` provides attributes indicating the agent responsible for some aspect of the text, the markup or something asserted by the markup, and the degree of certainty associated with it.

- `@cert` (certainty) signifies the degree of certainty associated with the intervention or interpretation.
- `@resp` (responsible party) indicates the agency responsible for the intervention or interpretation, for example an editor or transcriber.

`att.global.source` provides attributes used by elements to point to an external source.

- `@source` specifies the source from which some aspect of this element is drawn.

`att.editLike` provides attributes describing the nature of an encoded scholarly intervention or interpretation of any kind.

- `@evidence` indicates the nature of the evidence supporting the reliability or accuracy of the intervention or interpretation.

`att.dimensions` provides attributes for describing the size of physical objects.

- `@unit` names the unit used for the measurement
- `@quantity` specifies the length in the units specified
- `@extent` indicates the size of the object concerned using a project-specific vocabulary combining quantity and units in a single string of words.
- `@precision` characterizes the precision of the values specified by the other attributes.
- `@scope` where the measurement summarizes more than one observation, specifies the applicability of this measurement.

Many of the elements discussed here can be used in two ways. Their primary purpose is to indicate that the text encoded as the element's content represents an editorial intervention (or non-intervention) of a specific kind, indicated by the element itself. However, pairs or other meaningful groupings of such elements can also be supplied, wrapped within a special purpose <choice> element:

`<choice>` (choice) groups a number of alternative encodings for the same point in a text.

This element enables the encoder to represent for example a text in its 'original' uncorrected and unaltered form, alongside the same text in one or more 'edited' forms. This usage permits software to switch automatically between one 'view' of a text and another, so that (for example) a stylesheet may be set to display either the text in its original form or after the application of editorial interventions of particular kinds.

Elements which can be combined in this way constitute the `model.choicePart` class. The default members of this class are <sic>, <corr>, <reg>, <orig>, <unclear>, <supplied>, <abbr>, <expan>, <ex>, <am> and <seg>; some of their functions and usage are described further below.

Three categories of editorial intervention are discussed in this section:

- indication or correction of apparent errors
- indication or regularization of variant, irregular, non-standard, or eccentric forms
- editorial additions, suppressions, and omissions

A more extended treatment of the use of these tags in transcriptional and editorial work is given in chapter *12. Representation of Primary Sources*.

### 3.5.1 Apparent Errors

When the copy text is manifestly faulty, an encoder or transcriber may elect simply to correct it without comment, although for scholarly purposes it will often be more generally useful to record both the correction and the original state of the text. The elements described here enable all three approaches, and allows the last to be done in such a way as make it easy for software to present either the original or the correction.

`<sic>` (Latin for thus or so) contains text reproduced although apparently incorrect or inaccurate.

`<corr>` (correction) contains the correct form of a passage apparently erroneous in the copy text.

The following examples show alternative treatment of the same material. The copy text reads:

> Another property of computer-assisted historical research is that data modelling must permit any one textual feature or part of a textual feature to be a part of more than one information model and to allow the researcher to draw on several such models simultaneously, for example, to select from a machine-readable text those marginal comments which indicate that the date's mentioned in the main body of the text are incorrect.

An encoder may choose to correct the typographic error, either silently or with an indication that a correction has been made, as follows:

```
… marginal comments which indicate that the <corr>dates</corr>
mentioned in the main body of the text are incorrect.
```

Alternatively, the encoder may simply record the typographic error without correcting it, either without comment or with a <sic> element to indicate the error is not a transcription error in the encoding:

```
… marginal comments which indicate that the <sic>date's</sic>
mentioned in the main body of the text are incorrect.
```

If the encoder elects both to record the original source text and to provide a correction for the sake of word-search and other programs, both <sic> and <corr> are used, wrapped in a <choice>:

```
… marginal comments which indicate that the
<choice>
  <corr>dates</corr>
  <sic>date's</sic>
</choice> mentioned in the main body of the text are
incorrect.
```

The <sic> and <corr> elements can appear in either order.

If it is desired to indicate the person or edition responsible for the emendation, this might be done as follows:

```
… marginal comments which indicate that the
<choice>
  <corr  resp="#msm">dates</corr>
  <sic>date's</sic>
</choice> mentioned in the main body of the text are
incorrect.

<!-- within the header for this document ... -->
<respStmt  xml:id="msm">
  <resp>editor</resp>
  <name>C.M. Sperberg-McQueen</name>
</respStmt>
```

Here the *resp* attribute has been used to indicate responsibility for the correction. Its value (#msm) is an example of the *pointer* values discussed in section *3.7. Simple Links and Cross-References*; in this case, it points to a <respStmt> element within the TEI header, but any element might be indicated in this way, including for example a <name> element, or (if the module described in *14. Names, Dates, People, and Places* has been included) a <person> element. The *resp* attribute is available for all elements which are members of the `att.global.responsibility` class. The same class makes available a *cert* attribute, which may be used to indicate the degree of editorial confidence in a particular correction, as in the following example:

```
An <choice>
  <corr  cert="high">Autumn</corr>
  <sic>Antony</sic>
</choice> it was,
That grew the more by reaping
```

See further the discussion in section *12.3.1.3. Correction and Conjecture*.

Where, as here, the correction takes the form of adding text not otherwise present in the text being encoded, the encoder should use the <corr> element. Where the correction is present in the text being encoded, and consists of some combination of visible additions and deletions, the elements <add> or <del> should be used: see further section *3.5.3. Additions, Deletions, and Omissions* below. Where the correction takes the form of addition of material not present in the original because of physical damage or illegibility, the <supplied> element may be used. Where the 'correction' is simply a matter of expanding an abbreviation the <ex> element may be used. These and other elements to support the detailed encoding of authorial or scribal interventions of this kind are all provided by the module described in chapter *12. Representation of Primary Sources*.

## 3.5.2 Regularization and Normalization

When the source text makes extensive use of variant forms or non-standard spellings, it may be desirable for a number of reasons to *regularize* it: that is, to provide 'standard' or 'regularized' forms equivalent to the non-standard forms.[48]

As with other such changes to the copy text, the changes may be made silently (in which case the TEI header should specify the types of silent changes made) or may be explicitly marked using the following elements:

`<reg>` (regularization) contains a reading which has been regularized or normalized in some sense.

`<orig>` (original form) contains a reading which is marked as following the original, rather than being normalized or corrected.

`<choice>` (choice) groups a number of alternative encodings for the same point in a text.

Typical applications for these elements include the production of editions intended for student or lay readers, linguistic research in which spelling or usage variation is not the main question at issue, production of spelling dictionaries, etc.

Consider this 16th-century text:

> how godly a dede it is to overthrowe so wicked a race the world may judge: for my part I thinke there canot be a greater sacryfice to God.

An encoder may choose to preserve the original spelling of this text, but simply flag it as nonstandard by using the <orig> element with no attributes specified, as follows:

```
<p>...how godly a <orig>dede</orig> it is to
<orig>overthrowe</orig> so wicked a race the
world may judge: for my part I <orig>thinke</orig>
there <orig>canot</orig> be a greater
<orig>sacryfice</orig> to God</p>
```

Alternatively, the encoder may simply indicate that certain words have been modernized by using the <reg> element with no attributes specified, as follows:

```
<p>...how godly a
<reg>deed</reg> it is to <reg>overthrow</reg> so wicked a race the
world may judge: for my part I <reg>think</reg>
there <reg>cannot</reg> be a greater
<reg>sacrifice</reg> to God.</p>
```

[48] In some contexts, the term *regularization* has a narrower and more specific significance than that proposed here: the <reg> element may be used for any kind of regularization, including normalization, standardization, and modernization.

Alternatively, the encoder may elect to record both old and new spellings, so that (for example) the same electronic text may serve as the basis of an old- or new-spelling edition:

```
<p>...how godly a <choice>
    <orig>dede</orig>
    <reg>deed</reg>
  </choice> it is to
<choice>
    <orig>overthrowe</orig>
    <reg>overthrow</reg>
  </choice> so wicked a race the
world may judge: for my part I <choice>
    <orig>thinke</orig>
    <reg>think</reg>
  </choice>
there <choice>
    <orig>canot</orig>
    <reg>cannot</reg>
  </choice> be a greater
<choice>
    <orig>sacryfice</orig>
    <reg>sacrifice</reg>
  </choice> to God.</p>
```

As elsewhere, the *resp* attribute may be used to specify the agency responsible for the regularization.

### 3.5.3 Additions, Deletions, and Omissions

The following elements are used to indicate when words or phrases have been omitted from, added to, or marked for deletion from, a text. Like the other editorial elements, they allow for a wide range of editorial practices:

`<gap>` (gap) indicates a point where material has been omitted in a transcription, whether for editorial reasons described in the TEI header, as part of sampling practice, or because the material is illegible, invisible, or inaudible.

`@reason` (reason) gives the reason for omission.

`<ellipsis>` (deliberately marked omission) indicates a purposeful marking in the source document signalling that content has been omitted, and may also supply or describe the omitted content.

`<unclear>` (unclear) contains a word, phrase, or passage which cannot be transcribed with certainty because it is illegible or inaudible in the source.

`@reason` indicates why the material is hard to transcribe.

`<add>` (addition) contains letters, words, or phrases inserted in the source text by an author, scribe, or a previous annotator or corrector.

`<del>` (deletion) contains a letter, word, or passage deleted, marked as deleted, or otherwise indicated as superfluous or spurious in the copy text by an author, scribe, or a previous annotator or corrector.

Encoders may choose to omit parts of the copy text for reasons ranging from illegibility of the source or impossibility of transcribing it, to editorial policy, e.g. a systematic exclusion of poetry or prose from an encoding. The full details of the policy decisions concerned should be documented in the TEI header (see section *2.3. The Encoding Description*). In these situations — where copy text is not transcribed due to editorial policy or because it is impossible to do so — the <gap> element may be used to indicate the omission. Attributes on the <gap> element optionally provide further information about the reason for the omission, its extent, and the person or agency responsible for it, as in the following examples:

```
<gap  reason="illegible"  unit="word"
  quantity="2"  resp="#editor04"/>
```

```
<gap  reason="overwriting illegible"
  extent="several characters"/>
```

Note that the extent of the gap may be marked precisely using attributes *unit* and *quantity*, or more descriptively using the *extent* attribute. Other, more detailed, options are also available for representing dimensions of any kind; see further *11.3.4. Dimensions*.

The <desc> element may be used to supply a description of the material omitted, where that is considered useful:

```
<gap  reason="sampling"  quantity="120"
  unit="lines">
  <desc>irrelevant commentary</desc>
</gap>
```

```
… Their arrangement with respect to Jupiter and to each other was as follows:
<gap  reason="sampling"  extent="restOfPage">
  <desc>astrological figure</desc>
</gap>
That is, there were two stars on the easterly side and one to the
west; …
```

The <ellipsis> element provides a means to encode the deliberate, marked omission of content in a source document. Within <ellipsis>, a required <metamark> is used to represent the characters or markings in the source that indicate the presence of omitted material. Encoders may optionally provide a <desc> element to describe the omission, as well as a <supplied> element to record the content omitted from the source at this point. The <ellipsis> element could be applied to 'ditto' marks when they indicate repetition of the same content, as in this example:

```
<table  rows="3"  cols="2">
  <head>Tent supplies</head>
  <row>
    <cell>
      <name>nylon tent</name>
    </cell>
    <cell>
      <num>1</num>
    </cell>
  </row>
  <row>
    <cell>
      <name>steel stakes</name>
    </cell>
    <cell>
      <num>6</num>
    </cell>
  </row>
  <row>
    <cell>
      <name>fiberglass poles</name>
    </cell>
    <cell>
      <ellipsis>
        <metamark  function="ellipsis">"</metamark>
```

```
      <supplied>
        <num>6</num>
      </supplied>
    </ellipsis>
  </cell>
 </row>
</table>
```

The <ellipsis> element may also apply to literary passages that omit the full names of places or people, leaving their identities to the imagination of the reader, as in the next example.

```
<p>I leave Dublin to-morrow for M<ellipsis>
   <metamark>——</metamark>
 </ellipsis> house. It is situated in the county of <ellipsis>
   <metamark>——</metamark>
 </ellipsis>, on the north-west coast of <placeName>Connaught</placeName>, which I am
told is the classic ground of <placeName>Ireland</placeName>.</p>
```

A manuscript of song lyrics provides a more complex example of <ellipsis> to encode a marked omission that begins in the middle of a verse line. Our encoding of the passage shows the use of <ellipsis> with <supplied> to

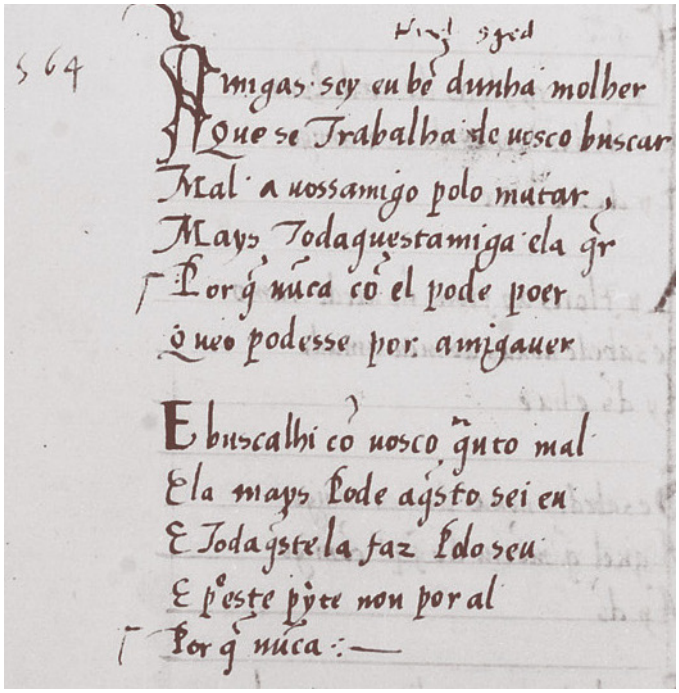



Figure 5: A manuscript from Cancioneiro da Biblioteca Nacional [1525-1526] in Biblioteca Nacional de Portugal, representing song lyrics and the beginning of a chorus. On the last line visible in this image, a marking of three dots and a line, '∴—', indicates that the rest of the chorus should follow after this point.

contain the omitted content.

```
<lg>
 <l>Amıgas sey eu be<ex>n</ex> dunha molher</l>
```

```
  <l>Que se trabalha de uosco buscar</l>
  <l>Mal a uossamigo polo matar</l>
  <l>Mays todaquestamiga ela q<ex>ue</ex>r</l>
  <metamark  function="signalChorus">ℾ</metamark>
  <l>Por q<ex>ue</ex> nu<ex>n</ex>ca co<ex>n</ex> el pode poer</l>
  <l>Queo podesse por amigauer</l>
</lg>
<lg>
  <l>E buscalhi co<ex>n</ex> uosco q<ex>ua</ex>nto mal</l>
  <l>Ela mays pode aq<ex>ue</ex>sto sei eu</l>
  <l>E todaq<ex>ue</ex>stela faz polo seu</l>
  <l>E p<ex>or</ex>este p<unclear>ỳ</unclear>te non por al</l>
  <lb/>
  <metamark  function="signalChorus">ℾ</metamark>
  <l  rend="nobreak"  part="I">Por q<ex>ue</ex> nu<ex>n</ex>ca</l>
  <ellipsis>
    <metamark  rend="nobreak"
      function="ellipsis">·:—</metamark>
    <supplied>
      <l  rend="nobreak"  part="F">co<ex>n</ex> el pode poer</l>
      <l>Queo podesse por amigauer</l>
    </supplied>
  </ellipsis>
</lg>
```

The <add> and <del> elements may be used to record where words or phrases have been added or deleted in the copy text. They are not appropriate where longer passages have been added or deleted, which span several elements; for these, the elements <addSpan> and <delSpan> described in chapter *12.3.1.4. Additions and Deletions* should be used.

Additions to a text may be recorded for a number of reasons. Sometimes they are marked in a distinctive way in the source text, for example by brackets or insertion above the line (*supralinear* insertion), as in the following example, taken from a 19th century manuscript:

```
The story I am going to relate is true as to its main facts,
and as to the consequences <add  place="above">of
these facts</add> from which this tale takes its title.
```

The <add> element should not be used to mark editorial changes, such as supplying a word omitted by mistake from the source text or a passage present in another version. In these cases, either the <corr> or <supplied> tags should be used, as discussed above in section *3.5.1. Apparent Errors*, and in section *12.3.1.3. Correction and Conjecture*, respectively.

The <unclear> element is used to mark passages in the original which cannot be read with confidence, or about which the transcriber is uncertain for other reasons, as for example when transcribing a partially inaudible or illegible source. Its *reason* and *resp* attributes are used, as with the <gap> element, to indicate the cause of uncertainty and the person responsible for the conjectured reading.

For example:

```
<l>And where the sandy mountain Fenwick scald</l>
<l>
  <unclear  reason="ink blot">The</unclear> sea between
yet hence his pray'r prevail'd
</l>
```

or from a spoken text:

```
<p>... and then <unclear  reason="passingTruck">marbled queen</unclear>...</p>
```

Where the material affected is entirely illegible or inaudible, the <gap> element discussed above should be used in preference.

The <del> element is used to mark material which is deleted in the source but which can still be read with some degree of confidence, as opposed to material which has been omitted by the encoder or transcriber either because it is entirely illegible or for some other reason. This is of particular importance in transcribing manuscript material, though deletion is also found in printed texts, sometimes for humorous purposes:

```
<l>One day I will sojourn to your shores</l>
<l>I live in the middle of England</l>
<l>But!</l>
<l>Norway! My soul resides in your watery
<del  rend="overstrike">fiords fyords fiiords</del>
</l>
<l>Inlets.</l>
```

The *rend* attribute may be used to distinguish different methods of deletion in manuscript or typescript material, as in this line from the typescript of Eliot's *Waste Land*:

```
<l>
  <del  rend="overtyped">Mein</del> Frisch <del  rend="overstrike">schwebt</del> weht
der Wind
</l>
```

Deletion in manuscript or typescript is often associated with addition:

```
<l>
  <del  rend="overstrike">Inviolable</del>
  <add  place="below">Inexplicable</add>
splendour of Corinthian white and gold
</l>
```

The <subst> element discussed in *12.3.1.5. Substitutions* provides a way of grouping additions and deletions of this kind.

The <del> element should not be used where the deletion is such that material cannot be read with confidence, or read at all, or where the material has been omitted by the transcriber or editor for some other reason. Where the material deleted cannot be read with confidence, the <unclear> tag should be used with the *reason* attribute indicating that the difficulty of transcription is due to deletion. Where material has been omitted by the transcriber or editor, this may be indicated by use of the <gap> element. A deletion in which some parts may be read but not others may thus be represented by one or more <gap> elements intermingled with text, all contained by a <del> element. Text supplied or marked as unneccessary by an editor should be marked with the <supplied> and <surplus> elements (discussed in *12.3.1.7. Text Omitted from or Supplied in the Transcription*) rather than <add> and <del>. These two sets of elements allow the encoder to distinguish editorial changes from those visible in the source text.

## 3.6 Names, Numbers, Dates, Abbreviations, and Addresses

This section describes a number of textual features which it is often convenient to distinguish from their surrounding text. Names, dates, and numbers are likely to be of particular importance to the scholar treating a text as source for a database; distinguishing such items from the surrounding text is however equally important to the scholar primarily interested in lexis.

The treatment of these textual features proposed here is not intended to be exhaustive: fuller treatments for names, numbers, measures, and dates are provided in the names and dates module (see chapter *14. Names, Dates, People, and Places*); more detailed treatment of abbreviations is provided by the transcription module (see section *12.3.1.2. Abbreviation and Expansion*).

### 3.6.1 Referring Strings

A *referring string* is a phrase which refers to some person, place, object, etc. Two elements are provided to mark such strings:

`<rs>` (referencing string) contains a general purpose name or referring string.

`<name>` (name, proper noun) contains a proper noun or noun phrase.

Both the <name> and <rs> elements are members of the `att.typed` class, from which they inherit the following attributes:

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

- `@type` characterizes the element in some sense, using any convenient classification scheme or typology.
- `@subtype` (subtype) provides a sub-categorization of the element, if needed.

which may be used to further categorize the kind of object referred to.

Examples include:

```
<p>
  <q>My dear
  <rs  type="person">Mr. Bennet</rs>
  </q>, said his lady to
him one day, <q>have you heard that <rs  type="place">Netherfield Park</rs> is let at
last?</q>
</p>
```

```
<p>Collectors of water-rents were appointed by the
<rs  type="org">Watering Committee</rs>.
They were paid a commission not exceeding four per
cent, and gave bond.</p>
```

```
<p>It being one of the principles of the
<rs  type="org">Circumlocution Office</rs> never, on any
account whatsoever, to give a straightforward answer,
<rs  type="person">Mr Barnacle</rs> said, <q>Possibly.</q>
</p>
```

As the following example shows, the <rs> element may be used for any reference to a person, place, etc., not only to references in the form of a proper noun or noun phrase.

```
<p>
  <q>My dear <rs  type="person">Mr. Bennet</rs>
  </q>, said
<rs  type="person">his lady</rs> to him one day ...
</p>
```

The <name> element by contrast is provided for the special case of referencing strings which consist only of proper nouns; it may be used synonymously with the <rs> element, or nested within it if a referring string contains a mixture of common and proper nouns. The following example shows an alternative way of encoding the short sentence from *Pride and Prejudice* quoted above:

```
<p>
  <q>My dear <name  type="person">Mr. Bennet</name>,</q> said <rs  type="person">his
lady</rs> to him one day,
<q>have you heard that <name  type="place">Netherfield Park</name> is let at last?</q>
</p>
```

As the following example shows, a proper name may be nested within a referring string:

```
<rs>His Excellency the Life President, <name>Ngwazi Dr H. Kamuzu Banda</name>
</rs>
```

Simply tagging something as a name is generally not enough to enable automatic processing of personal names into the canonical forms usually required for reference purposes. The name as it appears in the text may be inconsistently spelled, partial, or vague. Moreover, name prefixes such as *van* or *de la* may or may not be included as part of the reference form of a name, depending on the language and country of origin of the bearer.

Two issues arise in this context: firstly, there may be a need to encode a regularized form of a name, distinct from the actual form in the source to hand; secondly, there may be a need to identify the particular person, place, etc. referred to by the name, irrespective of whether the name itself is normalized or not. The element <reg>, introduced in *3.5.2. Regularization and Normalization* is provided for the former purpose; the attributes *key* or *ref* for the latter.

The *key* and *ref* attributes are common to all members of the `att.canonical` class and are defined as follows:

`att.canonical` provides attributes that can be used to associate a representation such as a name or title with canonical information about the object being named or referenced.

- `@key` provides an externally-defined means of identifying the entity (or entities) being named, using a coded value of some kind.
- `@ref` (reference) provides an explicit means of locating a full definition or identity for the entity being named by means of one or more URIs.

A very useful application for them is as a means of gathering together all references to the same individual or location scattered throughout a document:

```
<q>My dear
<rs  key="BENM1"  type="person">Mr. Bennet</rs>,</q> said
<rs  key="BENM2"  type="person">his lady</rs> to him one day,

<q>have you heard that
<rs  key="NETP1"  type="place">Netherfield Park</rs> is let at
last?</q>
```

```
<p>
  <name  key="VOM1"  type="person">Mme. de Volanges</name>
marie <rs  key="VOM2">sa fille</rs>:
c'est encore un secret;
mais elle m'en a fait part hier.
</p>
```

The value of the *key* attribute may be an unexpanded code, as in the examples above, with no particular significance. More usually however, it will be an externally defined code of some kind, as provided by a standard reference source.

```
<p>
  <name  key="LHR"  type="airport">Heathrow</name>
</p>
```

The standard reference source should be documented, for example using a <taxonomy> element in the TEI header.

The *ref* attribute can be used to point directly to some other resource providing more information about the entity named by the element, such as an authority record in a database, an encylopaedia entry, another element in the same or a different document etc.

```
<p>
  <name  ref="http://en.wikipedia.org/wiki/Heathrow_airport"
    type="airport">Heathrow</name>
</p>
```

This use should be distinguished from the use of a nested <reg> (regularization) element to provide the standard form of a referring string, as in this example:

```
<p>My personal life during
the administration of <rs  key="POJA1"  type="person">Col. Polk
    (<reg>Polk, James K.</reg>)</rs> has but poorly compensated me for the
suspended enjoyments and pursuits of private and professional
spheres</p>
```

No particular syntax is proposed for the values of the *key* attribute, since its form will depend entirely on practice within a given project. For the same reason, this attribute is not recommended in data interchange, since there is no way of ensuring that the values used by one project are distinct from those used by another. In such a situation, a preferable approach for magic tokens which follows standard practice on the Web is to use a *ref* attribute whose value is a tag URI as defined in RFC 4151. For example:

```
<p>
  <name  ref="tag:projectname.org,2012:VOM1"
    type="person">Mme. de Volanges</name> marie <rs  ref="tag:theworksoflaclos.org,2012:VOM2">sa
fille</rs>: c'est encore un secret;
mais elle m'en a fait part hier.
</p>
```

The inclusion of the domain name of the party responsible for tagging (`theworksoflaclos.org`), as specified in RFC 4151, helps ensure uniqueness of magic token values across TEI encoding projects, allowing for improved interchange of TEI documents.

The <choice> element discussed in *3.5. Simple Editorial Changes* may be used if it is desired to record both a normalized form of a name and the name used in the source being encoded:

```
<p>
  <name  ref="tag:projectname.org,2012:WADLM1"
    type="person">
    <choice>
      <orig>Walter de la Mare</orig>
      <reg>de la Mare, Walter</reg>
    </choice>
  </name>
was born at <name  ref="tag:projectname.org,2012:Ch1"
```

```
   type="place">Charlton</name>, in
<name  ref="tag:projectname.org,2012:KT1"
   type="county">Kent</name>, in 1873.
</p>
```

The <index> element discussed in *3.9.2. Index Entries* may be more appropriate if the function of the regularization is to provide a consistent index:

```
<p>
  <name  type="place">Montaillou</name> is not a large parish.
At the time of the events which led to
<name  type="person">Fournier<index>
     <term>Benedict XII, Pope of Avignon (Jacques Fournier)</term>
   </index>
  </name>'s
investigations, the local population consisted of between 200 and 250 inhabitants.
</p>
```

Although adequate for many simple applications, these methods have two inconveniences: if the name occurs many times, then its regularized form would be repeated many times; and the burden of additional XML markup in the body of the text may be inconvenient to maintain and complex to process. For applications such as onomastics, relating to persons or places named rather than the name itself, or wherever a detailed analysis of the component parts of a name is needed, the specialized elements described in chapter *14. Names, Dates, People, and Places* or the analytical tools described in chapter *19. Feature Structures* should be used.

### 3.6.2 Addresses

These Guidelines propose the following elements to distinguish postal and electronic addresses:

`<address>` (address) contains a postal address, for example of a publisher, an organization, or an individual.

`<email>` (electronic mail address) contains an email address identifying a location to which email messages can be delivered.

These two elements constitute the class of `model.addressLike` elements; for other kinds of address this class may be extended by adding new elements if necessary.

These Guidelines provide no particular means for encoding the substructure of an email address (for example, distinguishing the local part from the domain part), nor of distinguishing personal email addresses from generic or fictitious ones.

```
<email>info@tei-c.org</email>
```

The simplest way of encoding a postal address is to regard it as a series of distinct lines, just as they might be written on an envelope. The following element supports this view:

`<addrLine>` (address line) contains one line of a postal address.

Here is an example of a postal address encoded using this approach:

```
<address>
  <addrLine>110 Southmoor Road,</addrLine>
  <addrLine>Oxford OX2 6RB,</addrLine>
  <addrLine>UK</addrLine>
</address>
```

Alternatively, an address may be encoded as a structure of more semantically rich elements. The class `model.addrPart` element class identifies a number of such possible components:

`<street>` contains a full street address including any name or number identifying a building as well as the name of the street or route on which it is located.

`<name>` (name, proper noun) contains a proper noun or noun phrase.

`<postCode>` (postal code) contains a numerical or alphanumeric code used as part of a postal address to simplify sorting or delivery of mail.

`<postBox>` (postal box or post office box) contains a number or other identifier for some postal delivery point other than a street address.

`model.nameLike` groups elements which name or refer to a person, place, or organization.

| | |
|---|---|
| model.nameLike.agent | groups elements which contain names of individuals or corporate bodies. |
| model.offsetLike | groups elements which can appear only as part of a place name. |
| model.persNamePart | groups elements which form part of a personal name. |
| model.placeStateLike | groups elements which describe changing states of a place. |

| | |
|---|---|
| eventName | (name of an event) contains a proper noun or noun phrase used to refer to an event. |
| idno | (identifier) supplies any form of identifier used to identify some object, such as a bibliographic item, a person, a title, an organization, etc. in a standardized way. |
| lang | (language name) contains the name of a language mentioned in etymological or other linguistic discussion. |
| objectName | (name of an object) contains a proper noun or noun phrase used to refer to an object. |
| rs | (referencing string) contains a general purpose name or referring string. |

`model.persNamePart` groups elements which form part of a personal name.

| | |
|---|---|
| addName | (additional name) contains an additional name component, such as a nickname, epithet, or alias, or any other descriptive phrase used within a personal name. |
| forename | (forename) contains a forename, given or baptismal name. |
| genName | (generational name component) contains a name component used to distinguish otherwise similar names on the basis of the relative ages or generations of the persons named. |
| nameLink | (name link) contains a connecting phrase or link used within a name but not regarded as part of it, such as *van der* or *of*. |
| persPronouns | (personal pronouns) indicates the personal pronouns used, or assumed to be used, by the individual being described. |
| roleName | (role name) contains a name component which indicates that the referent has a particular role or position in society, such as an official title or rank. |
| surname | (surname) contains a family (inherited) name, as opposed to a given, baptismal, or nick name. |

`model.placeNamePart` groups elements which form part of a place name.

| | |
|---|---|
| bloc | (bloc) contains the name of a geo-political unit consisting of two or more nation states or countries. |

| | |
|---|---|
| country | (country) contains the name of a geo-political unit, such as a nation, country, colony, or commonwealth, larger than or administratively superior to a region and smaller than a bloc. |
| district | (district) contains the name of any kind of subdivision of a settlement, such as a parish, ward, or other administrative or geographic unit. |
| geogName | (geographical name) identifies a name associated with some geographical feature such as Windrush Valley or Mount Sinai. |
| placeName | (place name) contains an absolute or relative place name. |
| region | (region) contains the name of an administrative unit such as a state, province, or county, larger than a settlement, but smaller than a country. |
| settlement | (settlement) contains the name of a settlement such as a city, town, or village identified as a single geo-political or administrative unit. |

Any number of elements from the `model.addrPart` class may appear within an address and in any order. None of them is required.

Where code letters are commonly used in addresses (for example, to identify regions or countries) a useful practice is to supply the full name of the region or country as the content of the element, but to supply the abbreviatory code as the value of the global *n* attribute, so that (for example) an application preparing formatted labels can readily find the required information. Other components of addresses may be represented using the general-purpose <name> element or (when the additional module for names and dates is included) the more specialized elements provided for that purpose.

Using just the elements defined by the core module, the above address could thus be represented as follows:

```
<address>
  <street>110 Southmoor Road</street>
  <name  type="city">Oxford</name>
  <postCode>OX2 6RB</postCode>
  <name  type="country">United Kingdom</name>
</address>
```

The order of elements within an address is highly culture-specific, and is therefore unconstrained:

```
<address>
  <name  type="org">Università di Bologna</name>
  <name  type="country">Italy</name>
  <postCode>40126</postCode>
  <name  type="city">Bologna</name>
  <street>via Marsala 24</street>
</address>
```

A telephone number (normally outside of the <address> element) might be tagged with an <addrLine> and <ref> that points to the phone number using the `tel` URI scheme:

```
<addrLine>
  <ref  target="tel:+1-201-555-0123">(201) 555 0123</ref>
</addrLine>
```

For further discussion of ways of regularizing the names of places, see section *3.6. Names, Numbers, Dates, Abbreviations, and Addresses*. A full postal address may also include the name of the addressee, tagged as above using the general purpose <name> element.

When a schema includes the names and dates module discussed in chapter *14. Names, Dates, People, and Places*, a large number of more specific elements such as <country> or <settlement> will be available from the class `model.addrPart`. The above example might then be encoded as follows:

```
<address>
  <street>110 Southmoor Road</street>
  <settlement>Oxford</settlement>
  <postCode>OX2 6RB</postCode>
  <country>United Kingdom</country>
</address>
```

### 3.6.3 Numbers and Measures

This section describes elements provided for the simple encoding of numbers and measurements and gives some indication of circumstances in which this may usefully be done. The following phrase level elements are provided for this purpose:

`<num>` (number) contains a number, written in any form.

- `@type` indicates the type of numeric value.
- `@value` supplies the value of the number in standard form.

`<measure>` (measure) contains a word or phrase referring to some quantity of an object or commodity, usually comprising a number, a unit, and a commodity name.

- `@type` specifies the type of measurement in any convenient typology.

`<measureGrp>` (measure group) contains a group of dimensional specifications which relate to the same object, for example the height and width of a manuscript page.

Like names or abbreviations, numbers can occur virtually anywhere in a text. Numbers are special in that they can be written with either letters or digits (*twenty-one*, *xxi*, and *21*) and their presentation is language-dependent (e.g. English *5th* becomes Greek *5.*; English *123,456.78* equals French *123.456,78*).

For many kinds of application, e.g. natural-language processing or machine translation, numbers are not regarded as 'lexical' in the same way as other parts of a text. For these and other applications, the <num> element provides a convenient method of distinguishing numbers from the surrounding text. For other kinds of application, numbers are only useful if normalized: here the <num> element is useful precisely because it provides a standardized way of representing a numerical value.

For example:

```
<num value="33">xxxiii</num>
<num type="cardinal" value="21">twenty-one</num>
<num type="percentage" value="10">ten percent</num>
<num type="percentage" value="10">10%</num>
<num type="ordinal" value="5">5th</num>
```

```
<num type="fraction" value="0.5">one half</num>
<num type="fraction" value="0.5">1/2</num>
```

Sometimes it may be desired to mark something as numerical which cannot be accurately normalized, for example an expression such as *dozens*; less frequently the number may be recognisable linguistically as such but may use a notation with which the encoder is unfamiliar. To help in these situations, the <num> element may also bear either or both of the following attributes from the `att.ranging` class:

`att.ranging` provides attributes for describing numerical ranges.

- `@atLeast` gives a minimum estimated value for the approximate measurement.
- `@atMost` gives a maximum estimated value for the approximate measurement.

In its fullest form, a *measure* consists of a number, a phrase expressing units of measure and a phrase expressing the commodity being measured, though not all of these components need be present in every case. It may be helpful to distinguish measures from surrounding text for two reasons. Firstly, a measure may be expressed using

a particular notation or system of abbreviations which the encoder does not wish to regard as lexical. Secondly, a quantitative application may wish to distinguish and normalize the internal components of a measure, in order to perform calculations on them.

Consider, as an example of the first case, the following list of Celia's charms, in which the encoder has chosen to make explicit the measurements:

```
<div n="2">
  <list type="gloss">
    <label>Age</label>
    <item>Unimportant</item>
    <label>Head</label>
    <item>Small and round</item>
    <label>Eyes</label>
    <item>Green</item>
    <label>Complexion</label>
    <item>White</item>
    <label>Hair</label>
    <item>yellow</item>
    <label>Features</label>
    <item>Mobile</item>
    <label>Neck</label>
    <item>
      <measure>13¾"</measure>
    </item>
    <label>Upper arm</label>
    <item>
      <measure>11"</measure>
    </item>
<!--...-->
  </list>
<!-- ... -->
</div>
```

In the same way, it may be convenient to mark representations of currency which might otherwise be misinterpreted as lexical:

```
<p>...the sum of
<measure type="currency">12s 6d</measure>...</p>
```

In general, normalization of a measure will require specification of one or more of its three parts: the quantity, the units, and possibly also the commodity being measured. This is accomplished by supplying values for the three attributes *quantity*, *unit*, and *commodity*, which are supplied by the `att.measurement` class:

`att.measurement` provides attributes to represent a regularized or normalized measurement.

- `@quantity` (quantity) specifies the number of the specified units that comprise the measurement
- `@unit` (unit) indicates the units used for the measurement, usually using the standard symbol for the desired units.
- `@commodity` (commodity) indicates the substance that is being measured

With these attributes, the measurement of Celia's neck may be specified in a normalized form:

```
<measure quantity="13.75" unit="in">13¾"</measure>
```

Such techniques are particularly useful when representing historical data such as inventories:

```
<list>
  <item>
    <measure type="volume" quantity="2"
      unit="bag" commodity="hops">ii bags hops</measure>
  </item>
  <item>
    <measure type="volume" quantity="6"
      unit="truss" commodity="cloth">six trusses Woolen and linen goods</measure>
  </item>
  <item>
    <measure type="weight" quantity="5"
      unit="ton" commodity="coal">5 tonnes coale</measure>
  </item>
</list>
```

In addition to normalization, the attributes of the `att.ranging` class may be used to supply the mininum or maximum values of a range, or of an approximate measurement:

`att.ranging` provides attributes for describing numerical ranges.

- `@min` where the measurement summarizes more than one observation or a range, supplies the minimum value observed.
- `@max` where the measurement summarizes more than one observation or a range, supplies the maximum value observed.
- `@atLeast` gives a minimum estimated value for the approximate measurement.
- `@atMost` gives a maximum estimated value for the approximate measurement.

```
<p>In a cadaver I have found it
<lb/>holding <measure unit="mL" commodity="blood"
   min="85.2">more than three ounces</measure>.
[...] Then
<lb/>we may suppose in man that a single heart beat
<lb/>would force out either
<measure unit="mL" commodity="blood"
   max="13.2" min="3.7">a half ounce, three drams,
  <lb/>or even one dram of blood</measure>, which because of the
<lb/>valvular block could not flow back that way into
<lb/>the heart.</p>
```

The <measureGrp> element is provided as a means of grouping several related measurements together, either because the measurement involves several dimensions (for example height and width) or to avoid the need to repeat all the normalizing attributes:

```
<measureGrp type="volume" unit="in">
  <measure type="height" quantity="14">xiv</measure>
  <measure type="width" quantity="5">v</measure>
  <measure type="depth" quantity="10">x</measure>
</measureGrp>
```

The <unit> element may be applied when units of measurement require more detailed encoding about the system they belong to or the what kind of phenomenon they designate. The <unit> element may carry the *unit* attribute to indicate a standard value, as well as other optional attributes for indicating type, language, and other distinguishing characteristics.

`<unit>` contains a symbol, a word or a phrase referring to a unit of measurement in any kind of formal or informal system.

`att.measurement` provides attributes to represent a regularized or normalized measurement.
- `@unit` (unit) indicates the units used for the measurement, usually using the standard symbol for the desired units.

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

`att.global` provides attributes common to all elements in the TEI encoding scheme.

A <measure> element may contain a combination of <num> and <unit> elements:

```
<measure type="list">
  <num>1</num>, <num>2</num>, <num>5</num>, <num>7</num>
  <unit type="length" unit="mm">millimètres</unit>
</measure>
```

The unit element may also be nested to indicate a complex unit and its component parts, for example, to indicate that rate combines space and time:

```
<p>Light travels at <num value="3E10">3×10^10</num>
  <unit type="rate" unit="cm/s">
    <unit type="space">cm</unit> per <unit type="time">second</unit>
  </unit>.</p>
```

### 3.6.4 Dates and Times

Dates and times, like numbers, can appear in widely varying culture- and language-dependent forms, and can pose similar problems in automatic language processing. Such elements constitute the `model.dateLike` class, of which the default members are:

`<date>` (date) contains a date in any format.

`<time>` (time) contains a phrase defining a time of day in any format.

These elements have some additional attributes by virtue of being members of the `att.datable` and `att.duration` classes which, in turn, are members of the `att.datable.w3c` and `att.duration.w3c` classes. In particular, the *when* and *calendar* attributes will be discussed here:

`att.datable.w3c` provides attributes for normalization of elements that contain datable events conforming to the W3C XML Schema Part 2: Datatypes Second Edition.
- `@when` supplies the value of the date or time in a standard form, e.g. yyyy-mm-dd.

`att.calendarSystem` provides attributes for indicating calendar systems to which a date belongs.
- `@calendar` indicates one or more systems or calendars to which the date represented by the content of this element belongs.

Dates can occur virtually anywhere in a text, but in some contexts (e.g. bibliographic citations) their encoding is recommended or required rather than optional. Times can also appear anywhere but encoding these is more generally optional.

Partial dates or times (e.g. *1990*, *September 1990*, *twelvish*) can be expressed in the *when* attribute by simply omitting a part of the value supplied. Imprecise dates or times (for example *early August*, *some time after ten and before twelve*) may be expressed as date or time ranges.

These mechanisms are useful primarily for fully specified dates or times known with certainty. If component parts of dates or times are to be marked up, or if a more complex analysis of the meaning of a temporal expression is required, the techniques described in chapter *14. Names, Dates, People, and Places* should be used in preference to the simple method outlined here.

Where the certainty (i.e. reliability) of the date or time is in question, the encoder should record this fact using the mechanisms discussed in chapter *22. Certainty, Precision, and Responsibility*. The same chapter also discusses various methods of recording the precision of numerical or temporal assertions.

The *when* attribute is a useful way of normalizing or disambiguating dates and times which can appear in many formats, as the following examples show:

```
<date  when="1980-02-12">12/2/1980</date>
```

```
Given on the <date  when="1977-06-12">Twelfth Day of June
in the Year of Our Lord One Thousand Nine Hundred and
Seventy-seven of the Republic the Two Hundredth and first
and of the University the Eighty-Sixth.</date>
```

The *when* attribute always supplies a normalized representation of the date given as content of the <date> element. The format used should be a valid W3C schema datatype.[49] Some typical examples follow:

```
<date  when="2001">The year 2001</date>
<date  when="2001-09">September 2001</date>
<date  when="2001-09-11">11 Sep 01</date>
<date  when="--09-11">9/11</date>
<date  when="--09">September</date>
<date  when="---11">Eleventh of the month</date>
<time  when="08:48:00">8:48</time>
<date  when="2001-09-11T12:48:00">Sept 11th, 12 minutes before 9 am</date>
```

Note in the last example the use of a normalized representation for the date string which includes a time: this example could thus equally well be tagged using the <time> element.

The following examples demonstrate the use of the <date> element to mark a period of time:

```
<p>Those five years —
<date  from="1918"  to="1923">1918 to 1923</date>
— had been, he suspected,
somehow very important.</p>
```

```
<p>The Eddic poems are preserved in a unique
manuscript (Codex Regius 2365) from
<date  notBefore="1250"  notAfter="1300">the second half of the
    thirteenth century</date>, and <title>Hervarar
    saga</title> dates from <date  when="1300">around 1300</date>.</p>
```

The *calendar* attribute may be used to specify a date in any calendar system; if the *when* attribute is also supplied, it should specify the equivalent date in the Gregorian calendar.

[49] The datatypes are taken from the W3C Recommendation XML Schema Part 2: Datatypes Second Edition. The permitted datatypes are:

- date
- gYear
- gMonth
- gDay
- gYearMonth
- gMonthDay
- time
- dateTime

There is one exception: these Guidelines permit a time to be expressed as only a number of hours, or as a number of hours and minutes, as per ISO 8601:2004 section 4.2.2.3 and 4.3.3. The W3C `time` and `dateTime` datatypes require that the minutes and seconds be included in the normalized value if they are to be correctly processed for example when sorting.

### 3.6.5 Abbreviations and Their Expansions

It is sometimes desirable to mark abbreviations in the copy text, whether to trigger special processing for them, to provide the full form of the word or phrase abbreviated, or to allow for different possible expansions of the abbreviation. Abbreviations may be transcribed as they stand, or expanded; they may be left unmarked, or marked using these tags:

`<abbr>` (abbreviation) contains an abbreviation of any sort.

`<expan>` (expansion) contains the expansion of an abbreviation.

The <abbr> element is useful as a means of distinguishing semi-lexical items such as acronyms or jargon:

```
We can sum up the above discussion as follows:
the identity of a <abbr>CC</abbr> is defined by that calibration of values which
motivates the elements of its <abbr>GSP</abbr>; ...
```

```
Every manufacturer of <abbr>3GL</abbr> or <abbr>4GL</abbr>
languages is currently nailing on <abbr>OOP</abbr> extensions.
```

The *type* attribute may be used to distinguish types of abbreviation by their function:

```
<abbr  type="title">Dr.</abbr>
<abbr  type="initial">M.</abbr> Deegan is
the Director of the <abbr  type="acronym">CTI</abbr> Centre for Textual Studies.
```

Abbreviations such as *Dr. M.* above may be treated as two abbreviations, as above, or as one:

```
<abbr>Dr. M.</abbr> Deegan is
the Director of the <abbr>CTI</abbr> Centre for Textual Studies.
```

The <expan> element may be used simply to record that an abbreviation has been silently expanded by the encoder, perhaps for reasons of house style or editorial policy. It should always include the whole of an abbreviated phrase or word. More usually however this will be combined with the <abbr> element inside a <choice> element to record both the abbreviation and its expansion:

```
 the
<choice>
  <expan>World Wide Web Consortium</expan>
  <abbr>W3C</abbr>
</choice>
```

Nested abbreviations may also be handled in this way:

```
<choice>
  <abbr>RELAX NG</abbr>
  <expan>regular
     language for <choice>
      <abbr>XML</abbr>
      <expan>extensible markup
            language</expan>
    </choice>, next
     generation</expan>
</choice>
```

Abbreviation is a particularly important feature of manuscript and other source materials, the transcription of which needs more detailed treatment than is possible using these simple elements. A more detailed set of recommendations is discussed in *12.3.1. Altered, Corrected, and Erroneous Texts*, which includes additional elements made available for the purpose by the `transcr` module.

## 3.7 Simple Links and Cross-References

Cross-references or links between one location in a document and one or more other locations, either in the same or different XML documents, may be encoded using the elements <ptr> and <ref>, as discussed in this section. These elements both 'point' from one location in a document, the place that the element itself appears, to another (or to several), specified by means of a *target* attribute, supplied by the `att.pointing` class:

`att.pointing` provides a set of attributes used by all elements which point to other elements by means of one or more URI references.

`@target` specifies the destination of the reference by supplying one or more URI References.

Linkages of several other kinds are also provided for in these guidelines; see further chapter *17. Linking, Segmentation, and Alignment*.

The value of the *target* attribute, wherever it appears, provides a way of pointing to some other element using a method standardized by the W3C consortium, and known as the *XPointer* mechanism. This permits a range of complexity, from the very simple (a reference to the value of the target element's *xml:id* attribute) to the more complex usage of a full URI with embedded XPointers. For example, the source of the following paragraph looks something like this:

```
<p>...
The complete XPointer specification is managed by the W3C<note  place="foot">
   <ptr  target="https://www.w3.org/TR/xptr-framework/"/>,
  <ptr  target="https://www.w3.org/TR/xptr-element/"/>,
  <ptr  target="https://www.w3.org/TR/xptr-xmlns/"/>, and
  <ptr  target="https://www.w3.org/TR/xptr-xpointer/"/>
  </note>;
for a discussion of TEI schemes for XPointer, see
<ptr  target="#eSATS"/>.</p>
<!--... -->
<div  xml:id="eSATS">
<!--... -->
</div>
```

Alternatively, if no explicit link is to be encoded, but it is simply required to mark the phrase as a cross-reference, the <ref> element may be used without a *target* attribute.

For an introduction to the use of links in general, see *17. Linking, Segmentation, and Alignment*. The complete XPointer specification is managed by the W3C[50]; for a discussion of TEI schemes for XPointer, see *17.2.4. TEI XPointer Schemes*.

`<ptr>` (pointer) defines a pointer to another location.

`<ref>` (reference) defines a reference to another location, possibly modified by additional text or comment.

The elements <ptr> and <ref> are the default members of the phrase-level model class `model.ptrLike`. As members of the classes `att.pointing`, `att.typed`, `att.cReferencing`, and `att.internetMedia` they also carry the following attributes:

`att.pointing` provides a set of attributes used by all elements which point to other elements by means of one or more URI references.

`@target` specifies the destination of the reference by supplying one or more URI References.

`@evaluate` (evaluate) specifies the intended meaning when the target of a pointer is itself a pointer.

[50] `https://www.w3.org/TR/xptr-framework/`, `https://www.w3.org/TR/xptr-element/`, `https://www.w3.org/TR/xptr-xmlns/`, and `https://www.w3.org/TR/xptr-xpointer/`

`att.cReferencing` provides attributes that may be used to supply a *canonical reference* as a means of identifying the target of a pointer.

- `@cRef` (canonical reference) specifies the destination of the pointer by supplying a canonical reference expressed using the scheme defined in a <refsDecl> element in the TEI header.

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

- `@type` characterizes the element in some sense, using any convenient classification scheme or typology.
- `@subtype` (subtype) provides a sub-categorization of the element, if needed.

`att.internetMedia` provides attributes for specifying the type of a computer resource using a standard taxonomy.

- `@mimeType` (MIME media type) specifies the applicable multimedia internet mail extension (MIME) media type.

The two elements may be used in the same way; the difference between them is simply that while the <ptr> element is empty, the <ref> element may contain phrases specifying, or describing more exactly, the target of a cross-reference, which form the content of the element. Since its content thus serves as a human-readable pointer, in the simplest case a <ref> element need not identify its target in any other way. For example:

```
See <ref>section 12 on page 34</ref>.
```

More usually, it will be desirable to identify the target of the cross-reference using either the *target* or the *cRef* attribute, so that processing software can access it directly, for example to implement a linkage, to generate an appropriate reference, or to give an error message if it cannot be found. Assuming that section 12 in the previous example has been tagged

```
<div1 xml:id="SEC12">
<!-- ... -->
</div1>
```

then the same cross-reference might more exactly be encoded as

```
See especially <ref target="#SEC12">section 12 on page 34</ref>.
```

If the cross-reference itself is to be generated according to a fixed pattern, or if no text is to appear in the body of the cross-reference, the <ptr> element would be used as follows:

```
See in particular <ptr target="#SEC12"/>.
```

The *cRef* attribute may be used to express the target of a cross reference using some canonical referencing scheme, such as those typically used for ancient texts. In this case, the referencing scheme must be defined using the <cRefPattern> or <citeStructure> elements discussed below (*3.11.4. Declaring Reference Systems*); the definition these provide may be used to translate the value of the *cRef* attribute into a conventional pointer value, such as one that might be supplied by the *target* attribute. It is an error to supply both *cRef* and *target* values.

When the *target* attribute is used, a cross reference may point to any number of locations simultaneously, simply by giving more than one identifier as the value of its *target* attribute. This may be particularly useful where an analytic index is to be encoded, as in the following example:

```
<list>
  <item>Saints aid rejected in mel. <ptr  target="#p299"/>
  </item>
  <item>Sallets censured <ptr  target="#p143 #p144"/>
  </item>
  <item>Sanguine mel. signs <ptr  target="#p263"/>
  </item>
  <item>Scilla or sea onyon, a purger of mel. <ptr  target="#p442"/>
  </item>
</list>
```

Here the targets of the cross-references are simply page numbers; it is assumed that corresponding elements with identifiers p299, p143, etc. have been provided in the body of the text:

```
<pb  xml:id="p143"/>
...
<pb  xml:id="p144"/>
...
<pb  xml:id="p263"/>
...
<pb  xml:id="p299"/>
...
<pb  xml:id="p442"/>
...
```

A similar method may be used to link annotations on a text with the sigla used to encode their points of attachment in a text. For example:

```
annotated text <ref  target="#a51"  type="noteAnchor">51</ref>
<!-- ... -->
<note  xml:id="a51"  type="footnote">text of annotation</note>
```

The *type* attribute may be used, as elsewhere, to categorize the cross-reference according to any system of importance to the encoder. If bibliographic references require special processing (e.g. in order to provide a consistent short-form reference), they might be tagged thus:

```
Similar forms, often called
<term  rend="ldquo rdquo">rewriting systems</term>, have a long history
among mathematicians, but the specific form of <ptr  target="#fig22"/>
was first studied extensively by Chomsky <ptr  type="bibliog"  target="#chom59"/>.

<!-- ... -->
<figure  xml:id="fig22">
  <graphic  url="fig22.jpg"/>
</figure>
<!-- elsewhere, in the bibliography -->
<bibl  xml:id="chom59">
<!-- citation for the book referenced above -->
</bibl>
```

The value bibliog for the *type* attribute on the second <ptr> element here might be used to indicate that the object being referenced here is a bibliographic entry rather than a simple cross-reference to an illustration, as is the first <ptr>. In either case, the value of the *target* attribute is a pointer to some other element.

The <ptr> and <ref> elements have many applications in addition to the simple cross-referencing facilities illustrated in this section. In conjunction with the analytic tools discussed in chapters *17. Linking, Segmentation, and Alignment*, *18. Simple Analytic Mechanisms*, and *19. Feature Structures*, they may be used to link analyses of a text to their object, to combine corresponding segments of a text, or to align segments of a text with a temporal or other axis or with each other.

Where the *target* attribute of <ptr> or <ref> points to an external resource available on the network, the *mimeType* attribute may be used to specify the mime type of that resource; this may be important for to enable appropriate processing. For example:

```
<p>The current version of the TEI Guidelines source code
is available in the TEI GitHub Repository; <ref  target="https://github.com/TEIC/TEI/blob/dev/P5/S
   mimeType="application/tei+xml">guidelines-en.xml</ref>
is the root document used to create the English version
of these Guidelines.</p>
```

## 3.8 Lists

The following elements are provided for the encoding of lists, their constituent items, and the labels or headings associated with them:

`<list>` (list) contains any sequence of items organized as a list.

`<item>` (item) contains one component of a list.

`<label>` (label) contains any label or heading used to identify part of a text, typically but not exclusively in a list or glossary.

`<head>` (heading) contains any type of heading, for example the title of a section, or the heading of a list, glossary, manuscript description, etc.

`<headLabel>` (heading for list labels) contains the heading for the label or term column in a glossary list or similar structured list.

`<headItem>` (heading for list items) contains the heading for the item or gloss column in a glossary list or similar structured list.

The <list> element may be used to mark any kind of *list*: numbered, lettered, bulleted, or unmarked. Lists formatted as such in the copy text should in general be encoded using this element, with an appropriate value for the *rend* attribute. Suggested values for *rend* include:

- *bulleted* (items preceded by bullets or similar markings)
- *inline* (items rendered within continuous prose, with no line divisions)
- *numbered* (items preceded by numbers or letters)
- *simple* (items rendered as blocks, but with no bullet or number)

Some of these values may of course be combined; a list may be inline, but also be rendered with numbers. An example appears below. For more sophisticated and detailed description of list rendering, consider using the *style* attribute with Cascading Stylesheet properties and values, as described in the W3C's CSS Lists and Counters Module Level 3.

Each distinct item in the list should be encoded as a distinct <item> element. If the numbering or other identification for the items in a list is unremarkable and may be reconstructed by any processing program, no enumerator need be specified. If however an enumerator is retained in the encoded text, it may be supplied either by using the *n* attribute on the <item> element, or by using a <label> element. The following examples are thus equivalent:

```
I will add two facts, which have seldom occurred in
the composition of six, or even five quartos.
<list  rend="inline numbered">
  <label>(1)</label>
  <item>My first rough manuscript, without any
     intermediate copy, has been sent to the press.</item>
```

```
  <label>(2)</label>
  <item>Not a sheet has been seen by any human
     eyes, excepting those of the author and the printer:
     the faults and the merits are exclusively my own.</item>
</list>
```

```
I will add two facts, which have seldom occurred in
the composition of six, or even five quartos.
<list  rend="inline numbered">
  <item  n="1">My first rough manuscript, without any
     intermediate copy, has been sent to the press.</item>
  <item  n="2">Not a sheet has been seen by any human
     eyes, excepting those of the author and the printer:
     the faults and the merits are exclusively my own.</item>
</list>
```

The two styles may not be mixed in the same list: if one item is preceded by a label, all must be.

A list need not necessarily be displayed in list format. For example, the following is a reasonable encoding of a list which (in the original) is simply printed as a single paragraph:

```
On those remote pages it is written that animals are
divided into <list  rend="inline">
  <item  n="a">those that belong to the Emperor, </item>
  <item  n="b">embalmed ones, </item>
  <item  n="c">those that are trained, </item>
  <item  n="d">suckling pigs, </item>
  <item  n="e">mermaids, </item>
  <item  n="f">fabulous ones, </item>
  <item  n="g">stray dogs, </item>
  <item  n="h">those that are included in this classification, </item>
  <item  n="i">those that tremble as if they were mad, </item>
  <item  n="j">innumerable ones, </item>
  <item  n="k">those drawn with a very fine camel's-hair brush, </item>
  <item  n="l">others, </item>
  <item  n="m">those that have just broken a flower vase, </item>
  <item  n="n">those that resemble flies from a distance. </item>
</list>
```

A list may be given a heading or title, for which the <head> element should be used, as in the next example, which also demonstrates simple use of the <label> element to mark a tabular or glossary list in which each item is associated with a word or phrase rather than a numeric or alphabetic enumerator:

```
<list  type="gloss">
  <head>Report of the conduct and progress of Ernest Pontifex.
     Upper Vth form — half term ending Midsummer 1851</head>
  <label>Classics</label>
  <item>Idle listless and unimproving</item>
  <label>Mathematics</label>
  <item>ditto</item>
  <label>Divinity</label>
  <item>ditto</item>
  <label>Conduct in house</label>
```

```
  <item>Orderly</item>
  <label>General conduct</label>
  <item>Not satisfactory, on account of his great
     unpunctuality and inattention to duties</item>
</list>
```

In such a list, the individual items have internal structure. In complex cases, where list items contain many components, the list is better treated as a *table*, on which see chapter *15. Tables, Formulæ, Graphics, and Notated Music*. A particularly important instance of the simple two-column table is the 'glossary list', which should be marked by the tag <list type="gloss">. In such lists, each <label> element contains a term and each <item> its gloss; it is a semantic error for a list tagged with `type="gloss"` not to have labels. For example:

```
<list  type="gloss">
  <head>Unit Three — Vocabulary</head>
  <label  xml:lang="la">acerbus, -a, -um </label>
  <item>bitter, harsh</item>
  <label  xml:lang="la">ager, agrī, M. </label>
  <item>field</item>
  <label  xml:lang="la">audiō, īre,
     īvī, ītus </label>
  <item>hear, listen (to)</item>
  <label  xml:lang="la">bellum, -ī, N. </label>
  <item>war</item>
  <label  xml:lang="la">bonus, -a, -um </label>
  <item>good</item>
</list>
```

Additionally, the <term> and <gloss> elements discussed in section *3.4.1. Terms and Glosses* might be used to make explicit the role that each column in the glossary list has, as follows:

```
<list  type="gloss">
  <head>Unit Three — Vocabulary</head>
  <label>
    <term  xml:lang="la">acerbus, -a, -um</term>
  </label>
  <item>
    <gloss>bitter, harsh</gloss>
  </item>
  <label>
    <term  xml:lang="la">ager, agrī, M. </term>
  </label>
  <item>
    <gloss>field</gloss>
  </item>
  <label>
    <term  xml:lang="la">audiō, -īre, -īvī, -ītus</term>
  </label>
  <item>
    <gloss>hear, listen (to)</gloss>
  </item>
  <label>
    <term  xml:lang="la">bellum, -ī, N. </term>
  </label>
  <item>
```

```
    <gloss>war</gloss>
  </item>
  <label>
    <term  xml:lang="la">bonus, -a, -um</term>
  </label>
  <item>
    <gloss>good</gloss>
  </item>
</list>
```

Note in the above examples the use of the global *xml:lang* attribute to specify on the <label> (or <term>) element what language the term is from. For further discussion of the *xml:lang* attribute see section *1.3.1.1. Global Attributes*, and section *vi.1 Language Identification.* A more elaborate markup for this glossary would distinguish the headword forms from the grammatical information (principal parts and gender), perhaps using elements taken from *10. Dictionaries*.

In addition to the <head> element used to supply a title or heading for the whole list, headings for the two columns of a glossary-style list may be specified using the two special elements <headLabel> and <headItem>:

```
The simple, straightforward statement of an idea is
preferable to the use of a worn-out expression.
<list  type="gloss">
  <headLabel>TRITE</headLabel>
  <headItem>SIMPLE, STRAIGHTFORWARD</headItem>
  <label>bury the hatchet </label>
  <item>stop fighting, make peace</item>
  <label>at loose ends </label>
  <item>disorganized</item>
  <label>on speaking terms </label>
  <item>friendly</item>
  <label>fair and square </label>
  <item>completely honest</item>
  <label>at death's door </label>
  <item>near death</item>
</list>
```

The elements <label>, <head>, <headLabel>, and <headItem> may contain only phrase-level elements. The <item> element however may contain paragraphs or other 'chunks', including other lists. In this example, a glossary list contains two items, each of which is itself a simple list:

```
<list  type="gloss">
  <label>EVIL</label>
  <item>
    <list  rend="bulleted">
      <item>I am cast upon a horrible desolate island, void
            of all hope of recovery.</item>
      <item>I am singled out and separated as it were from
            all the world to be miserable.</item>
      <item>I am divided from mankind — a solitaire; one
            banished from human society.</item>
    </list>
  </item>
  <label>GOOD</label>
  <item>
    <list  rend="bulleted">
      <item>But I am alive; and not drowned, as all my
```

```
            ship's company were.</item>
        <item>But I am singled out, too, from all the ship's
            crew, to be spared from death...</item>
        <item>But I am not starved, and perishing on a barren place,
            affording no sustenances....</item>
      </list>
    </item>
  </list>
```

Lists of different types may be nested to arbitrary depths in this way.

## 3.9 Notes, Annotation, and Indexing

### 3.9.1 Notes and Simple Annotation

The following element is provided for the encoding of discursive notes, whether already present in the copy text or supplied by the encoder:

`<note>` (note) contains a note or annotation.

A note is any additional comment found in a text, marked in some way as being out of the main textual stream. All notes should be marked using the same tag, <note>, whether they appear as block notes in the main text area, at the foot of the page, at the end of the chapter or volume, in the margin, or in some other place.

Notes may be in a different hand or typeface, may be authorial or editorial, and may have been added later. Attributes may be used to specify these and other characteristics of notes, as detailed below.

A note is usually attached to a specific point or span within a text, which we term here its *point of attachment*. In conventional printed text, the point of attachment is represented by some siglum such as a star or cross, or a superscript digit.

When encoding such a text, it is conventional to replace this siglum by the content of the annotation, duly marked up with a <note> element. This may not always be possible for example with marginal notes, which may not be anchored to an exact location. For ease of processing, it may be adequate to position marginal notes before the relevant paragraph or other element. In printed texts, it is sometimes conventional to group notes together at the foot of the page on which their points of attachment appear. This practice is not generally recommended for TEI-encoded texts, since the pagination of a particular printed text is unlikely to be of structural significance. In some cases, however, it may be desirable to transcribe notes not at their point of attachment to the text but at their point of appearance, typically at the end of the volume, or the end of the chapter. In such cases, the *target* attribute of the <note> may be used to indicate the point of attachment. It is also possible to encode the point of attachment itself, using the <ptr> or <ref> element, pointing from that to the body of the <note> placed elsewhere.

In cases where the note is applied not to a point but to a span of text, not itself represented as a TEI element, the *target* attribute may use an appropriate pointer expression, for example using the `range()` function to specify the span of attachment.

For further discussion of pointing to points and spans in the text, see section *3.7. Simple Links and Cross-References*.

In the following example, the *type* attribute is used to categorise the note as a gloss:

```
<l>The self-same moment I could pray</l>
<l>And from my neck so free</l>
<l>The albatross fell off, and sank</l>
<l>Like lead into the sea.
<note  type="gloss"  place="margin">The spell begins to break</note>
</l>
```

As the <note> appears within an <l> element, we may infer that its point of attachment is in the margin adjacent to the line in question. In the following version of the same text, however, it may be inferred that the note applies to the whole of the stanza:

```
<lg>
  <l>The self-same moment I could pray</l>
  <l>And from my neck so free</l>
  <l>The albatross fell off, and sank</l>
  <l>Like lead into the sea.</l>
  <note  type="gloss"  place="margin">The spell begins to break</note>
</lg>
```

This type of annotation, very common in the early printed texts which Coleridge may be presumed to be imitating in this case, may also be regarded as providing a heading or descriptive label for the passage concerned. The encoder may therefore prefer to use the <label> element to represent it, as in the following case:

```
<lg>
  <l>The self-same moment I could pray</l>
  <l>And from my neck so free</l>
  <l>The albatross fell off, and sank</l>
  <l>Like lead into the sea.</l>
  <label  place="margin">The spell begins to break</label>
</lg>
```

In the following example, a note which appears at the foot of the page in the printed source is given at its point of attachment within the text. The global *n* attribute is used to indicate the note number:

```
Collections are ensembles of
distinct entities or objects of any sort.<note  n="1"  place="bottom">We
explain below why we use the uncommon term
<mentioned>collection</mentioned> instead of the expected
<mentioned>set</mentioned>. Our usage corresponds to the
<mentioned>aggregate</mentioned> of many mathematical writings and to
the sense of <mentioned>class</mentioned> found in older logical
writings.</note> The elements ...
```

In addition to transcribing notes already present in the copy text, researchers may wish to add their own notes or comments to it. The <note> element may be used for either purpose, but it will usually be advisable to distinguish the two categories. One way might be to use the *type* attribute shown above, categorizing notes as *authorial*, *editorial*, etc. Where notes derive from many sources, or where a more precise attribution is required, the *resp* attribute may be used to point to a definition of the person or other agency responsible for the content of the note.

As a simple example, an edition of the *Ancient Mariner* might include both Coleridge's original glosses and those of a modern commentator:

```
<lg>
<!-- ... -->
  <l>The self-same moment I could pray;
  <note  place="margin"  resp="#STC"
     type="gloss">The spell begins to break</note>
    <note  place="bottom"  resp="#JLL">The turning point of the poem...</note>
  </l>
</lg>
```

For this to be valid, the codes #JLL and #STC must point to some more information identifying the agency concerned. The syntax used is identical to that used for other cross-references, as discussed in *3.7. Simple Links and Cross-References*; thus in this case, the TEI header for this text might contain a title statement like the following:

```
<titleStmt>
  <title>The Rime of the Ancient Mariner: an annotated edition</title>
  <author  xml:id="STC">Samuel Taylor Coleridge</author>
  <editor  xml:id="JLL">John Livingston Lowes</editor>
</titleStmt>
```

When annotating the electronic text by means of analytic notes in some structured vocabulary, e.g. to specify the topics or themes of a text, the  and <interp> elements may be more effective than the free form <note> element; these elements are available when the module for simple analysis is selected (see section *18.3. Spans and Interpretations*).

#### 3.9.1.1 Encoding Grouped Notes

The following element is provided for the grouping of notes:

`<noteGrp>` (note group) contains a group of notes.

A text may have multiple alternative versions of the same note, such as the same annotation expressed in multiple languages, or both an extensive note and a short form for different audiences. In such cases multiple <note> elements may be grouped within a <noteGrp> element.

Typically, the <note> elements within a <noteGrp> would be differentiated by use of attributes such as *xml:lang* or *type*, while sharing the same point of attachment. This differentiation can be made either implicitly in case of inline notes, or explicitly via a *target* attribute, which may be specified on the <noteGrp> itself.

The simple example below demonstrates the grouping of a short and a full version of the same note, where the short version might be intended for use in contexts with space constraints.

```
<p>... reuerendos dominos archiepiscopum et canonicos Leopolienses
in duplicibus Quatuortemporibus
<noteGrp>
   <note  type="short">Quatuor Tempora, so called dry fast days.</note>
   <note  type="full">Quatuor Tempora, so called dry fast days (Wednesday, Friday, and
Saturday)
       falling on each of the quarters of the year. In the first quarter they were
called Cinerum
       (following Ash Wednesday), second Spiritus (following Pentecost), third Crucis
       (after the Exaltation of the Holy Cross, September 14th), and Luciae
       in the fourth (after the feast of St. Lucia, December 13th).
   </note>
  </noteGrp>
totaliter expediui.
</p>
```

### 3.9.2 Index Entries

The indexing of scholarly texts is a skilled activity, involving substantial amounts of human judgment and analysis. It should not therefore be assumed that simple searching and information retrieval software will be able to meet all the needs addressed by a well-crafted manual index, although it may complement them for example by providing free text search. The role of an index is to provide access via keywords and phrases which are not necessarily present in the text itself, but must be added by the skill of the indexer.

#### 3.9.2.1 Pre-existing Indexes

When encoding a pre-existing text, therefore, if such an index is present it may be advisable to retain it along with the text, rather than attempt to regenerate it automatically. Elements discussed elsewhere in these Guidelines may be used for this purpose. For example, the <div1> element or  element may be used to mark the section of the text containing the index and the <list> element might be used to mark the index itself, each entry being represented by an <item> element, possibly containing within it a series of <ptr> or <ref> elements, as follows:

```
<div  type="index">
<!--...-->
  <list  type="index">
    <item>Women, how cause of mel. <ref>193</ref>; their vanity in
        apparell taxed, <ref>527</ref>; their counterfeit tears
    <ref>547</ref>; their vices <ref>601</ref>, commended,
    <ref>624</ref>.</item>
    <item>Wormwood, good against mel. <ref>443</ref>
    </item>
    <item>World taxed, <ref>181</ref>
    </item>
    <item>Writers of the cure of mel. 295</item>
<!--...-->
  </list>
</div>
```

Note that this simple representation does not capture the nested structure of the first of these index entries. A more accurate representation might entail the use of nested lists like the following:

```
<item>Women,
<list>
    <item>how cause of mel. <ref>193</ref>;</item>
    <item>their vanity in apparell taxed, <ref>527</ref>;</item>
    <item>their counterfeit tears <ref>547</ref>;</item>
    <item>their vices
    <list>
        <item>
          <ref>601</ref>,</item>
        <item> commended, <ref>624</ref>.</item>
      </list>
    </item>
  </list>
</item>
```

The page references, encoded simply as <ref> elements above, might also include direct links to the appropriate location in the encoded text, using (for example) a target attribute to supply the identifier of an associated <pb> element:

```
<!-- in the text --><pb  xml:id="P624"/>
<!-- start of page 624 -->
<!-- in the index -->
<ref  target="#P624">624</ref>
```

For further discussion of this and alternative ways of encoding such links see the discussion in section *17. Linking, Segmentation, and Alignment.* Note that similar methods may also be used to encode a table of contents, as further exemplified in section *4.5. Front Matter.*

#### 3.9.2.2 Auto-generated Indexes

It can also be useful, however, to generate a new index from a machine-readable text, whether the text is being written for the first time with the tags here defined, or as an addition to a text transcribed from some other source. Depending on the complexity of the text and its subject matter, such an automatically-generated index may not in itself satisfy all the needs of scholarly users. However it can assist a professional indexer to construct a fully adequate index, which might then be post-edited into the digital text, marked-up along the lines already suggested for preserving pre-existing index material.

Indexes generally contain both references to specific pages or sections and references to page ranges or sequences. The same element is used in either case:

`<index>` (index entry) marks a location to be indexed for whatever purpose.

Like the <interp> element described in *18.3. Spans and Interpretations* this element may be used simply to provide descriptive or interpretive label of some kind for any location within a text, to be processed in any way by analytic software, but its main purpose is to facilitate the generation of an index for a printed version of the text. An <index> element may be placed anywhere within a text, between or within other elements. The headwords to be used when making up this index are given by the <term> elements within the <index> element. The location of the generated index might be specified by means of a processing instruction within the text, such as the following (the exact form of the PI is of course dependent on the application software in use):

```
<?tei indexplacement ?>
```

Alternatively, the special purpose <divGen> element might be used.

In the simplest case, a single headword is supplied by a <term> element contained by an <index> element:

```
<p>The students understand procedures for Arabic lemmatisation
<index>
   <term>Lemmatization, Arabic</term>
 </index>and are beginning to build parsers.</p>
```

The effect of this is to document an index entry for the term 'Lemmatization, Arabic', which when processed could reference the location of the original <index> element.

If the subject of Arabic lemmatization is treated at length in a text, then the index entry generated may need to reference a sequence of locations (e.g. page numbers). In such a case it will be necessary to identify the end of the relevant span of text as well as its starting point. This is most conveniently done by supplying an empty <anchor> element (as discussed in chapter *17. Linking, Segmentation, and Alignment*) at the appropriate point and pointing to it from the <index> element by means of its *spanTo* attribute, as in this example:

```
<p>We now turn to the
topic of Arabic lemmatisation
<index  spanTo="#ALAMEND">
   <term>Lemmatization, Arabic</term>
 </index> concerning which it is important to note [...]

<!-- much learned material omitted here -->
and now we can build our parser.<anchor  xml:id="ALAMEND"/>
</p>
```

This would generate the same index entries as the previous example, but the reference would be to the whole span of text between the location of the <index> element and the location of the element identified by the code `ALAMEND`, rather than a single point, and thus might (for example) include a sequence of page numbers.

Although the position of the <index> element in the text provides the target location that will be specified in the generated index entry, no part of the text itself is used to construct that entry. Index terms appearing in the entry come solely from the content of <term> elements, which consequently may have to repeat words or phrases from the text proper. This need not be done verbatim, thus giving scope for normalization of spelling (as in the example above) or other modifications which may assist generation of an index in a desired form or sequence.

Sometimes, for example when index terms are taken from a different language or consist of mathematical formulae or other expressions, even a normalized form of an index term may be insufficient for an application to order it exactly as desired. The *sortKey* attribute may be used to address this problem, as in the following example:

```
<p>The @ operator
<index>
   <term  sortKey="0000">@</term>
  </index> precedes an
attribute name</p>
```

Here, an entry for the symbol @ will appear in the index, but will be sorted alphabetically as if it were the string 0000. This technique is also useful when an index entry is to contain some non-Unicode character or glyph represented by the <g> element discussed in chapter *5. Characters, Glyphs, and Writing Modes*. In the following example, we assume that somewhere a definition for this glyph has been provided using the elements described in chapter *5. Characters, Glyphs, and Writing Modes*, and given the code PrinceGlyph:

```
<char  xml:id="PrinceGlyph">
<!-- definition of the glyph here -->
</char>
<p>The Artist formerly known as Prince <index>
   <term  sortKey="Prince">
     <g  ref="#PrinceGlyph"/>
   </term>
  </index>...</p>
```

Note that if no value is supplied for the sortKey attribute, a sorting application should always use the content of the <term> element as a sort key.

It is common practice to compile more than one index for a given text. A biography of a poet, for example, may offer an index of references to poems by the subject of the study, another index of works by other writers, an index of places or historical personages etc. The indexName attribute is used to assigning index terms and locations to one or more specific indexes:

```
<p>Sir John Ashford
<index  indexName="INDEX-PERSONS">
   <term>Ashford, John</term>
  </index> was,
coincidentally, born in
<index  indexName="INDEX-PLACES">
   <term>Ashford
       (Kent)</term>
  </index>Ashford...</p>
```

Multi-level indexing is particularly common in scholarly documents. For example, as well as entries such as *TEI*, or *markup*, an index may contain structured entries like *TEI, markup practices, index terms*, where a top level entry *TEI* is followed by a number of second-level subcategories, any or all of which may have a third-level list attached to them and so on. In order to reflect such a hierarchical index listing, <index> elements may be nested to the required depth. For example, suppose that we wish to make a structured index entry for 'lemmatisation' with subentries for 'Arabic', 'Sanskrit', etc. The example at the start of this section might then be encoded with nested <index> elements:

```
<p>The students understand procedures for Arabic lemmatisation
<index>
   <term>lemmatization</term>
   <index>
     <term>arabic</term>
   </index>
  </index>
...</p>
```

The index entry from Burton's *Anatomy of Melancholy* quoted above might be generated in a similar way. To generate such an entry, the body of the text might include, at page 193, an <index> element such as

```
<index>
  <term>Women</term>
  <index>
    <term>how cause of mel.</term>
  </index>
</index>
```

. Similarly, page 601 of the body text would include an <index> element like the following:

```
<index>
  <term>Women</term>
  <index>
    <term>their vices</term>
  </index>
</index>
```

while the <index> element at page 624 would have a structure like the following:

```
<index>
  <term>Women</term>
  <index>
    <term>their vices</term>
    <index>
      <term>commended</term>
    </index>
  </index>
</index>
```

When processing such <index> elements, the duplication required to make the structure explicit will normally be removed, so as to produce entries like those quoted above. However, this is not required by the encoding recommended here.

As noted above, either a processing instruction or a <divGen> element may be used to mark the place at which an index generated from <index> elements should be inserted into the output of a processing program; typically but not necessarily this will be at some point within the back matter of the document. If the <divGen> element is used, then the *type* attribute should be used to specify which kind of index is to be generated, and its value should correspond with that of the *indexName* attribute on the relevant <index> elements.

```
<back>
  <div  type="appendix">
    <head>Bibliography</head>
    <listBibl>
      <bibl> ... </bibl>
    </listBibl>
  </div>
  <divGen  n="Index Nominum"
    type="INDEX-NAMES"/>
  <divGen  n="Index Loci"  type="INDEX-PLACES"/>
</back>
```

As this example shows, the global *n* attribute may also be used to specify a name or identifier for the generated index itself in the usual way. Any additional headings etc. required for the generated index must be specified as content of the <divGen> element.

```
<back>
  <divGen  n="A1"  type="INDEX-NAMES">
    <head>An Index of Names</head>
  </divGen>
</back>
```

If a processing instruction is used, then these parameters for the generated index may be supplied in some other way.

One final feature frequently found in manually-created indexes to printed works cannot readily be encoded by the means provided here, namely cross-references internal to the index term listing. For example, if all references to the TEI in a text have been indexed using the index term *Text Encoding Initiative*, it may also be helpful to include an entry under the term *TEI* containing some text such as 'see Text Encoding Initiative'. Such internal cross-references must be added as part of the post-editing phase for an auto-generated index.

## 3.10 Graphics and Other Non-textual Components

Graphics, such as illustrations or diagrams, appear in many different kinds of text, and often with different purposes. Audio or video clips may also appear. In some cases, such media form an integral part of a text (indeed, some texts—comic books for example—may be almost entirely graphic); in others the graphic or video may be a kind of optional extra. In some cases, the text may be incomprehensible unless the media is included; in others, the presence of the media adds little to the sense of the work. It will therefore be a matter of encoding policy as to whether or how media found in a source text are transferred to a new encoded version of the same. In documents which are 'born digital', media such as graphics and other non-textual components may be particularly salient, but their inclusion in an archival form of the document concerned remains an editorial decision.

Considered as structural components, media may be anchored to a particular point in the text, or they may *float* either completely freely, or within some defined scope, such as a chapter or section. Time-based media such as audio or video may need to be synchronized with particular parts of a written text. Media of all kinds often contain associated text such as a heading or label. These Guidelines provide the following different elements to indicate their appearance within a text:

`<figure>` (figure) groups elements representing or containing graphic information such as an illustration, formula, or figure.

`<media>` indicates the location of any form of external media such as an audio or video clip etc.

`<graphic>` (graphic) indicates the location of a graphic or illustration, either forming part of a text, or providing an image of it.

`<binaryObject>` provides encoded binary data representing an inline graphic, audio, video or other object.

Media files may be encoded in a number of different ways:

- in some non-XML or binary format such as PNG, JPEG, MP3, MP4 etc.
- in an XML format such as SVG
- in a TEI XML format such as the notation for graphs and trees described in *20. Graphs, Networks, and Trees*

In the last two cases, the presence of the graphic will be indicated by an appropriate XML element, drawn from the SVG namespace in the second case, and its content will fully define the graphic to be produced. In the first case, however, one of the elements <graphic> or <media> is used to mark the presence of the graphic only and the visual content itself is stored outside the XML document at a location referenced by means of a *url* attribute. This attribute is provided by membership of these elements in the `att.resourced` class. Alternatively, if it is small, the media information may be embedded directly within the document using some suitable binary format such as Base64; in this case the <binaryObject> element may be used to contain it.

The elements <graphic>, <media>, and <binaryObject> are made available as members of the class `model.graphicLike` when this module is included in a schema. These elements are also members of the class `att.media`, from which they inherit the following attributes:

`att.internetMedia` provides attributes for specifying the type of a computer resource using a standard taxonomy.

- `@mimeType` (MIME media type) specifies the applicable multimedia internet mail extension (MIME) media type.

`att.media` provides attributes for specifying display and related properties of external media.

- `@width` Where the media are displayed, indicates the display width.
- `@height` Where the media are displayed, indicates the display height.
- `@scale` Where the media are displayed, indicates a scale factor to be applied when generating the desired display size.

For example, the following passage indicates that a copy of the image found in the source text may be recovered from the URL `zigzag2.png` and that this image is in PNG format:

```
<p>These were the four lines I moved in
through my first, second, third, and
fourth volumes. — In the fifth volume
I have been very good, — the precise
line I have described in it being this:
<graphic url="zigzag2.png"
   mimeType="image/png"/>
By which it appears, that except at the
curve, marked A. where I took a trip
to Navarre, — and the indented curve B.
which is the short airing when I was
there with the Lady Baussiere and her
page, — I have not taken the least frisk
...</p>
```

The media elements are phrase level elements which may be used anywhere that textual content is permitted, within but not between paragraphs or headings. In the following example, the encoder has decided to treat a specific printer's ornament as a heading:

```
<head>
  <graphic url="http://www.iath.virginia.edu/gants/Ornaments/Heads/hp-ral02.gif"/>
</head>
```

.

The <figure> element discussed in *15.4. Specific Elements for Graphic Images* provides additional capabilities, for example the ability to combine a number of images into a hierarchically organized structure or a block of images. The <figure> element carries a *type* attribute, which can be used to distinguish different kinds of graphic component within a single work, for example, maps as opposed to illustrations. It also provides the ability to associate an image with additional information such as a heading or a description.

## 3.11 Reference Systems

By *reference system* we mean the system by which names or references are associated with particular passages of a text (e.g. *Ps. 23:3* for the third verse of Psalm 23 or *Amores 2.10.7* for Ovid's *Amores*, book 2, poem 10, line 7). Such names make it possible to mark a place within a text and enable other readers to find it again. A reference system may be based on structural units (chapters, paragraphs, sentences; stanza and verse), typographic units (page and line numbers), or divisions created specifically for reference purposes (chapter and verse in Biblical texts). Where one exists, the traditional reference system for a text should be preserved in an electronic transcript of it, if only to make it easier to compare electronic and non-electronic versions of the text.

Reference systems may be recorded in TEI-encoded texts in any of the following ways:

- where a reference system exists, and is based on the same logical structure as that of the text's markup, the reference for a passage may be recorded as the value of the global *xml:id* or *n* attribute on an appropriate tag, or may be constructed by combining attribute values from several levels of tags, as described below in section *3.11.1. Using the xml:id and n Attributes*.
- where there is no pre-existing reference system, the global *xml:id* or *n* attributes may be used to construct one (e.g. collections and corpora created in electronic form), as described below in section *3.11.2. Creating New Reference Systems*.
- where a reference system exists which is not based on the same logical structure as that of the text's markup (for example, one based on the page and line numbers of particular editions of the text rather than on the structural divisions of it), any of a variety of methods for encoding the logical structure representing the reference system may be employed, as described in chapter *21. Non-hierarchical Structures*.
- where a reference system exists which does not correspond to any particular logical structure, or where the logical structure concerned is of no interest to the encoder except as a means of supporting the referencing system, then references may be encoded by means of <milestone> elements, which simply mark points in the text at which values in the reference system change, as described below in section *3.11.3. Milestone Elements*.

The specific method used to record traditional or new reference systems for a text should be declared in the TEI header, as further described in section *3.11.4. Declaring Reference Systems* and in section *17.2.5. Canonical References*.

When a text has no pre-existing associated reference system of any kind, these Guidelines recommend as a minimum that at least the page boundaries of the source text be marked using one of the methods outlined in this section. Retaining page boundaries in the markup is also recommended for texts which have a detailed reference system of their own. Line divisions in prose texts may be, but need not be, encoded.[51]

## 3.11.1 Using the *xml:id* and *n* Attributes

When traditional reference schemes represent a hierarchical structuring of the text which mirrors that of the marked-up document, the *n* attribute defined for all elements may be used to indicate the traditional identifier of the relevant structural units. The *n* attribute may also be used to record the numbering of sections or list items in the copy text if the copy-text numbering is important for some reason, for example because the numbers are out of sequence.

For example, a traditional reference to Ovid's *Amores* might be *Amores 2.10.7*—book 2, poem 10, line 7. Book, poem, and line are structural units of the work and will therefore be tagged in any case. (See chapter *6. Verse* for a discussion of structural units in verse collections.) In such cases, it is convenient to record traditional reference numbers of the structural units using the *n* attribute. The relevant tags for our example would be:

```
<div1 n="Amores" type="volume">
  <div2 n="1" type="book">
<!-- ... -->
  </div2>
  <div2 n="2" type="book">
    <div3 n="1" type="poem">
<!-- ... -->
    </div3>
    <div3 n="2" type="poem">
<!-- ... -->
    </div3>
<!-- ... -->
    <div3 n="10" type="poem">
      <l n="1"> ... </l>
      <l n="2"> ... </l>
<!-- ... -->
      <l n="7"> ... </l>
    </div3>
<!-- ... -->
```

[51] Many encoders find it convenient to retain the line divisions of the original during data entry, to simplify proofreading, but this may be done without inserting an element for the beginning of each line in the original.

```
  </div2>
<!-- ... -->
</div1>
```

One may also place the entire standard reference for each portion of the text into the appropriate value for the *n* attribute, though for obvious reasons this takes more space in the file:

```
<div1  n="Amores"  type="volume">
  <div2  n="Amores 1"  type="book">
<!-- ... -->
  </div2>
  <div2  n="Amores 2"  type="book">
    <div3  n="Amores 2.1"  type="poem">
<!-- ... -->
    </div3>
<!-- ... -->
    <div3  n="Amores 2.10"  type="poem">
<!-- ... -->
     <l  n="Amores 2.10.7"> ... </l>
<!-- ... -->
    </div3>
<!-- ... -->
  </div2>
<!-- ... -->
</div1>
```

If the names used by the traditional reference system can be formulated as identifiers, then the references can be given as values for the *xml:id* attribute; this requires that the reference be given without internal spaces, begin with a letter or underscore, and contain no characters other than letters, digits, hyphens, underscores, full stops, and the various combining and extender characters, as defined by the XML specification. Unlike values for the *n* attribute, values for the *xml:id* attribute must be unique throughout the document. Our example then looks like this:

```
<div1  n="Amores"  type="volume">
  <div2  xml:id="am.1"  type="book">
<!-- ... -->
  </div2>
  <div2  xml:id="am.2"  type="book">
    <div3  xml:id="am.2.1"  type="poem">
<!-- ... -->
    </div3>
<!-- ... -->
    <div3  xml:id="am.2.10"  type="poem">
<!-- ... -->
     <l  xml:id="am.2.10.7"> ... </l>
<!-- ... -->
    </div3>
<!-- ... -->
  </div2>
<!-- ... -->
</div1>
```

To document the usage and to allow automatic processing of these standard references, it is recommended that the TEI header be used to declare whether standard references are recorded in the *n* or *xml:id* attributes and which

elements may carry standard references or portions of them. For examples of declarations for the reference systems just shown, see section *3.11.4. Declaring Reference Systems*.

Using the *n* attribute one can specify only a single standard referencing system, a limitation not without problems, since some editions may define structural units differently and thus create alternative reference systems. For example, another edition of the *Amores* considers poem 10 a continuation of poem 9, and therefore would specify the same line as *Amores 2.9.31*. In order to record both of these reference systems one could employ any of a variety of methods discussed in chapter *21. Non-hierarchical Structures*.

## 3.11.2 Creating New Reference Systems

If a text has no canonical reference system of its own, a new custom reference system may be used.

The global attributes *n* and *xml:id* may be used to assign reference identifiers to segments of the text. Identifiers specified by either attribute apply to the entire element for which they are given. *xml:id* attributes must be unique within a single document, and *xml:id* values must begin with a letter. No such restrictions are made on the values of *n* attributes.

Determining a referencing system for a TEI encoding depends on many factors that may either be derived from textual structure, or influenced by extra-textual contingencies such as project and file management concerns. It is important, therefore, that the attribute used, the elements which can bear standard reference identifiers, and the method for constructing standard reference identifiers, should all be declared in the header as described in section *2.3.6. The Reference System Declaration*.

The Guidelines do not recommend one specific method for creating new referencing systems; however, the rest of this section lists some possibly useful strategies.

### 3.11.2.1 Referencing system derived from markup

A new referencing system may be derived from the structure of the electronic text, specifically from the markup of the text. As with any reference system intended for long-term use, it is important to see the reference as an established, unchanging point in the text. Should the text be revised or rearranged, the reference-system identifiers associated with any section of text must stay with that section of text, even if it means the reference numbers fall out of sequence. (A new reference system may always be created beside the old one if out-of-sequence numbers must be avoided.)

A convenient method of mechanically generating unique values for *xml:id* or *n* attributes based on the structure of the document is to construct, for each element, a *domain-style address* comprising a series of components separated by full stops, with one component for each level of the document hierarchy. Two methods may be used. In the *typed path* form of identifier, each component in the identifier takes the form of an element identifier, a hyphen, and a number, for example `p-2`. The element name specifies what type of element is to be sought, and the number specifies which occurrence of that element type is to be selected. (The hyphen and number may be omitted if there is only one element of the given type.) In the *untyped path* form of identifier, each component consists of a number, indicating which element in the sequence of nodes at each level is to be selected. To make the resulting identifier a valid XML identifier, it may need to be prefixed with an unchanging alphabetic letter.

Identifiers generated with these methods should use the <text> element as their starting point, rather than the <TEI> or <body> elements. The <TEI> element may be taken as a starting point only if identifiers need to be generated for the <teiHeader>, which is not usually the case; using the <body> element as a root would prevent assignment of identifiers for the front and back matter. The component corresponding to the root element can be omitted from identifiers, if no confusion will result. In collections and corpora, the component corresponding to the root may be replaced by the unique identifier assigned to the text or sample.

In the following example, each element within the <text> element has been given a typed-path identifier as its *xml:id* value, and an untyped-path identifier as its *n* value; the latter are prefixed with the string *AB*, which may be imagined to be the general identifier for this text.

```
<text xml:id="Text-1" n="AB">
  <front xml:id="Front" n="AB.1">
    <div xml:id="Front.div-1" n="AB.1.1">
      <p> ... </p>
    </div>
    <titlePage xml:id="Front.titlePage"
```

```
      n="AB.1.2">
      <titlePart> ... </titlePart>
    </titlePage>
    <div xml:id="Front.div-2" n="AB.1.3">
      <p> ... </p>
    </div>
  </front>
  <body xml:id="Body" n="AB.2">
    <p xml:id="Body.p-1" n="AB.2.1"> ... </p>
    <p xml:id="Body.p-2" n="AB.2.2"> ... </p>
    <div xml:id="Body.div-1" n="AB.2.3">
      <head xml:id="Body.div-1.head"
        n="AB.2.3.1"> ... </head>
      <p xml:id="Body.div-1.p-1" n="AB.2.3.2"> ... </p>
      <p xml:id="Body.div-1.p-2" n="AB.2.3.3"> ... </p>
    </div>
    <div xml:id="Body.div-2" n="AB.2.4">
      <head xml:id="Body.div-2.head"
        n="AB.2.4.1"> ... </head>
      <p xml:id="Body.div-2.p-1" n="AB.2.4.2"> ... </p>
      <p xml:id="Body.div-2.p-2" n="AB.2.4.3"> ... </p>
    </div>
  </body>
</text>
```

The typed and untyped path methods are convenient, but are in no way required for anyone creating a reference system.

If the *xml:id* attribute is used to record the reference identifiers generated, each value should record the entire path. If the *n* attribute is used, each value may record either the entire path or only the subpath from the parent element. The attribute used, the elements which can bear standard reference identifiers, and the method for constructing standard reference identifiers, should all be declared in the header as described in section *2.3.6. The Reference System Declaration.*

#### 3.11.2.2 Referencing systems based on project conventions

A reference system may be based on an agreed project-specific convention for *xml:id* attributes. Every convention will have strengths and weaknesses and it is left to encoders to make a decision that enables them to locate information in their TEI document.

Here are some examples of referencing systems that have been used in TEI project:

- **Title-based identifiers**: identifiers constructed with a number of characters from the main document title, followed by an incremental number. E.g. HOL001, HOL002, etc. using a fixed number of digits; or without fixed digits: HOL1, HOL2, etc.
- **Based on markup, with prefix**: identifiers constructed on the markup itself, as described in the previous section. To facilitate uniqueness in a corpus, each identifier may be prefixed with the identifier of the root <TEI> element. E.g. RootID-Body-p-1.
- **Opaque identifiers**: computed identifiers using either a randomized algorithm or a universally unique identifier (UUID) algorithm. Note that XSLT's function generate-id() only guarantees identifier unique to the document being processed.

XML well-formedness requires only that xml:id attributes be unique within a single document. However, it is also worth keeping in mind that for operating with referencing systems across a corpus of TEI files it is helpful (or even necessary in some circumstances) to have unique identifiers across the whole corpus.

Values of *xml:id* may be either populated computationally or manually. In the latter case, it is advisable to put

measures in place to avoid human error. Custom data types and Schematron rules may be defined in a customization ODD, and a check digit may be added to prevent unwanted changes. [52]

### 3.11.3 Milestone Elements

Where the desired reference system does not correspond to any particular structural hierarchy, or the document combines multiple structural hierarchies (as further discussed in *21. Non-hierarchical Structures*), simpler though less expressive methods may be necessary. In such cases the simplest solution may be just to mark up changes in the reference system where they occur, by using one or more of the following *milestone* elements:

`<milestone>` (milestone) marks a boundary point separating any kind of section of a text, typically but not necessarily indicating a point at which some part of a standard reference system changes, where the change is not represented by a structural element.

`<gb>` (gathering beginning) marks the beginning of a new gathering or quire in a transcribed codex.

`<pb>` (page beginning) marks the beginning of a new page in a paginated document.

`<lb>` (line beginning) marks the beginning of a topographic line in some edition or version of a text.

`<cb>` (column beginning) marks the beginning of a new column of a text on a multi-column page.

These elements simply mark the points in a text at which some category in a reference system changes. They have no content but subdivide the text into regions, rather in the same way as milestones mark points along a road, thus implicitly dividing it into segments. The elements <gb>, <pb>, <cb>, and <lb> are specialized types of milestone, marking gathering, page, column, and line boundaries respectively. The global *n* attribute is used in each case to provide a value for the particular unit associated with this milestone (for example, the page or line number). Since it is not structural, validation of a reference system based on <milestone>s cannot readily be checked by an XML parser, so it will be the responsibility of the encoder or the application software to ensure that they are given in the correct order.

Milestone elements are often used as a simple means of capturing the original appearance of an early printed text, which will rarely coincide exactly with structural units, but they are generally useful wherever a text has two or more competing structures. For example, many English novels were first published as serial works, individual parts of which do not always contain a whole number of chapters. An encoder might decide to represent the chapter-based structure using <div1> elements, with <milestone> elements to mark the points at which individual parts end; or the reverse. Thus, an encoding in which chapters are regarded as more important than parts might encode some work in which chapter three begins in part one and is concluded in part two as follows:

```
<text>
  <body>
    <milestone unit="part" n="1"/>
    <div1 n="1" type="chapter">
      <p>
<!-- ... -->
      </p>
    </div1>
    <div1 n="2" type="chapter">
      <p>
<!-- ... -->
      </p>
    </div1>
    <div1 n="3" type="chapter">
      <p>
<!-- ... -->
      </p>
      <milestone unit="part" n="2"/>
      <p>
<!-- ... -->
      </p>
```

[52] A check digit is computed from the value of an identifier and appended to the value itself. If the identifier is changed, the check digit would therefore invalidate it.

```
   </div1>
  </body>
</text>
```

An encoding of the same work in which parts are regarded as more important than chapters might begin as follows:

```
<text>
  <body>
    <div1  n="1"  type="part">
      <milestone  unit="chapter"  n="1"/>
      <p>
<!-- ... -->
      </p>
      <milestone  unit="chapter"  n="2"/>
      <p>
<!-- ... -->
      </p>
      <milestone  unit="chapter"  n="3"/>
      <p>
<!-- ... -->
      </p>
    </div1>
    <div1  n="2"  type="part">
      <p>
<!-- ... -->
      </p>
      <milestone  unit="chapter"  n="4"/>
      <p>
<!-- ... -->
      </p>
    </div1>
  </body>
</text>
```

Similarly, when tagging dramatic verse one may wish to privilege stanzas and lines over speeches and speakers, particularly where speeches cross line and line group boundaries. One might also wish to mark changes in narrative voice in a prose text. In either case, a milestone tag may be used to indicate change of speaker:

```
<lg>
  <milestone  unit="speaker"  n="Man"/>
  <l>Oh what is this I cannot see</l>
  <l>With icy hands gets a hold on me</l>
  <milestone  unit="speaker"  n="Death"/>
  <l>Oh I am Death, none can excel</l>
  <l>I open the doors of heaven and hell</l>
</lg>
```

Milestone tags also make it possible to record the reference systems used in a number of different editions of the same work. The reference system of any one edition can be recreated from a text in which all are marked by simply ignoring all elements that do not specify that edition on their *ed* attribute.

As a simple example, assuming that edition E1 of some collection of poems regards the first two poems as constituting the first book, while edition E2 regards the first poem as prefatory, a markup scheme like the following might be adopted:

```
<milestone ed="E1" unit="work"/>
<milestone ed="E2" unit="work"/>
<milestone ed="E1" unit="book"/>
<milestone ed="E1" unit="poem"/>
<milestone ed="E2" unit="poem"/>
<milestone ed="E2" unit="book"/>
<milestone ed="E1" unit="poem"/>
<milestone ed="E2" unit="poem"/>
```

In this case no *n* value is specified, since the numbers rise predictably and the application can keep a count from the start of the document, if desired.

The value of the *n* attribute may but need not include the identifiers used for any larger sections. That is, either of the following styles is legitimate:

```
<milestone ed="E1" unit="work" n="Amores"/>
<milestone ed="E1" unit="book" n="1"/>
<milestone ed="E1" unit="poem" n="1"/>
<milestone ed="E1" unit="poem" n="2"/>
<milestone ed="E1" unit="book" n="2"/>
```

or

```
<milestone ed="E1" unit="work" n="Amores"/>
<milestone ed="E1" unit="book" n="1"/>
<milestone ed="E1" unit="poem" n="1.1"/>
<milestone ed="E1" unit="poem" n="1.2"/>
<milestone ed="E1" unit="book" n="2"/>
```

When using <milestone> tags, line numbers may be supplied for every line or only periodically (every fifth, every tenth line). The latter may be simpler; the former is more reliable.

The style of numbering used in the values of *n* is unrestricted: for the example above, I.i, I.ii, and I.iii could have been used equally well if preferred. The special value unnumbered should be reserved for marking sections of text which fall outside the normal numbering system (e.g. chapter heads, poem numbers, titles, or speaker attributions in a verse drama).

By default, there are no constraints on the values supplied for the *ed* attribute. If it is felt appropriate to enforce such a restriction, the techniques described in *24.3. Customization* may be used, for example to specify that the attribute must specify one of a predefined set of values.

See below, section *3.11.4. Declaring Reference Systems*, for examples of declarations for the reference systems just shown.

Milestone elements may be used to mark any kind of shift in the properties associated with a piece of text, whether or not would normally be considered a reference system. For example, they may be used to mark changes in narrative voice in a prose text, or changes of speaker in a dramatic text, where these are not marked using structural elements such as <sp>, perhaps in order to avoid a clash of hierarchies.

As noted in *3.2.2. Hyphenation* above, milestone elements such as <lb> or <pb> represent whitespace and are therefore by default assumed to occur between orthographic tokens in the text, where these are not otherwise indicated. By default it is reasonable to assume that words are not broken across page or line boundaries, and that therefore a sequence such as

```
...sed imp<lb/>erator dixit...
```

should be tokenized as four words (*sed*, *imp*, *erator*, and *dixit*). The *break* attribute is provided to change the default assumption. To make explicit that *imperator* in the above example should be treated as a single word, a tagging such as the following is recommended:

```
...sed imp<lb  break="no"/>erator dixit...
```

Where hyphenation appears at the end of a topographic line, a column, or a page, the encoder may or may not choose to record the fact, either explicitly using an appropriate Unicode character, or descriptively for example by means of the *rend* attribute; see further *3.2.2. Hyphenation.*

### 3.11.4 Declaring Reference Systems

Whatever kind of reference system is used in an electronic text, it is recommended that the TEI header contain a description of its construction in the <refsDecl> element described in section *2.3.6. The Reference System Declaration.* As described there, the declaration may consist either of a formal declaration using the <cRefPattern> or <citeStructure> elements, or an informal description in prose. One of the former is recommended because unlike prose they can be processed by software.

The three examples given in section *3.11.1. Using the xml:id and n Attributes* would be declared as follows. The first example encodes the standard references for Ovid's *Amores* one level at a time, using the *n* attribute on the <div1>, <div2>, <div3>, and <l> tags. The header section for such an encoding should look something like this:

```
<encodingDesc>
  <refsDecl>
    <cRefPattern  matchPattern="([^]+) ([0-9]+)\.([0-9]+)\.([0-9]+)"
      replacementPattern="#xpath(//div1[@n='$1']/div2[@n='$2']/div3[@n='$3']/l[@n='$4']">
      <p>A canonical reference is assembled with
      <list>
          <item>the name of the <label>work</label>: the
          <att>n</att> of a <gi>div1</gi>,</item>
          <item>a space,</item>
          <item>the number of the <label>book</label>: the
          <att>n</att> of a child <gi>div2</gi>,</item>
          <item>a full stop</item>
          <item>the number of the <label>poem</label>: the
          <att>n</att> of a child <gi>div3</gi>,</item>
          <item>the line number: the <att>n</att> value of a
                   child <gi>l</gi>
          </item>
        </list>
      </p>
    </cRefPattern>
    <cRefPattern  matchPattern="([^]+) ([0-9]+)\.([0-9]+)"
      replacementPattern="#xpath(//div1[@n='$1']/div2[@n='$2']/div3[@n='$3']">
      <p>Same as above, but without the last component (full
            stop followed by the <gi>l</gi>'s <att>n</att>.</p>
    </cRefPattern>
    <cRefPattern  matchPattern="([^]+) ([0-9]+)"
      replacementPattern="#xpath(//div1[@n='$1']/div2[@n='$2']">
      <p>Same as above, but without the poem component (full
            stop followed by the <gi>div3</gi>'s <att>n</att>.</p>
    </cRefPattern>
  </refsDecl>
</encodingDesc>
```

The second example encodes the same reference system, again using the *n* attribute on the <div1>, <div2>, <div3>, and <l> tags, but giving the reference string in full on each tag. If canonical references are made only to lines, the reference system could be declared as follows:

```
<refsDecl>
  <cRefPattern  matchPattern="([^]+ [0-9]+\.[0-9]+\.[0-9]+)"
    replacementPattern="#xpath(//l[@n='$1'])"/>
</refsDecl>
```

Since the entire regular expression is enclosed as a parenthetical subgroup, the entire canonical reference string is sought as the value of the *n* attribute on an <l> element.

In order to handle references to poems as well as to individual lines, the declaration for the reference system must be more complicated:

```
<refsDecl>
  <cRefPattern  matchPattern="([^]+ [0-9]+\.[0-9]+\.[0-9]+)"
    replacementPattern="#xpath(//l[@n='$1'])"/>
  <cRefPattern  matchPattern="([^]+ [0-9]+\.[0-9]+)"
    replacementPattern="#xpath(//div2[@n='$1'])"/>
</refsDecl>
```

This declaration indicates that the entire reference string must be sought as the value of the *n* attribute on a <div1>, <div2>, <div3>, or <l> element.

The third example encodes the same reference system, this time giving the entire reference string as the value of the *xml:id* attribute on the relevant tags. The reference system declaration for such an encoding could be:

```
<refsDecl>
  <cRefPattern  matchPattern="(.*)"
    replacementPattern="#$1"/>
</refsDecl>
```

although in general there seems to be little advantage in this case: it is no more difficult to use a standard relative URI reference as the value of *target*.

In cases where a more complete formal declaration of text structure is desirable, for example in systems that will present the contents of a large TEI file in smaller chunks, the <citeStructure> element may be used. This method permits canonical references to be resolved and also allows them to be extracted, so that, for example, a list of resolvable citations may be generated from the document. The example from the *Amores* above could be implemented using <citeStructure> thus:

```
<refsDecl>
  <citeStructure  unit="work"  match="//div1"
    use="@n">
    <citeStructure  unit="book"  match="div2"
      use="@n"  delim=" ">
      <citeStructure  unit="poem"  match="div3"
        use="@n"  delim=".">
        <citeStructure  unit="line"  match="l"
          use="@n"  delim="."/>
      </citeStructure>
    </citeStructure>
  </citeStructure>
</refsDecl>
```

<citeStructure> also provides a method for attaching informational properties to units of structure, by means of the <citeData> element. The work, book, and poem divisions above might all have <head> elements which provide a title for the section. If we wish that information to be extractable, we can use citeData to specify where it is to be found:

```
<refsDecl>
  <citeStructure unit="work" match="//div1"
    use="@n">
    <citeData property="http://purl.org/dc/terms/title"
      use="head"/>
    <citeStructure unit="book" match="div2"
      use="@n" delim=" ">
      <citeData property="http://purl.org/dc/terms/title"
        use="head"/>
      <citeStructure unit="poem" match="div3"
        use="@n" delim=".">
        <citeData property="http://purl.org/dc/terms/title"
          use="head"/>
        <citeStructure unit="line" match="l"
          use="@n" delim="."/>
      </citeStructure>
    </citeStructure>
  </citeStructure>
</refsDecl>
```

The example above maps the head element to the Dublin Core property `title`. For convenience, property URIs may be abbreviated using <prefixDef>.

Reference systems recorded by means of milestone tags can also be declared; the following prose description could be used to declare the example given in section *3.11.3. Milestone Elements*.

```
<refsDecl>
  <p>Standard references to work, book, poem, and line may be
     constructed from the milestone tags in the text.</p>
</refsDecl>
```

Or in this way, using a formal declaration for this reference scheme derived from edition E1.

```
<refsDecl>
  <refState ed="E1" unit="work" delim=" "/>
  <refState ed="E1" unit="book" delim="."/>
  <refState ed="E1" unit="poem" delim=":"/>
  <refState ed="E1" unit="line"/>
</refsDecl>
```

## 3.12 Bibliographic Citations and References

Bibliographic references (that is, full descriptions of bibliographic items such as books, articles, films, broadcasts, songs, etc.) or pointers to them may appear at various places in a TEI text. They are required at several points within the TEI header's source description, as discussed in section *2.2.7. The Source Description*; they may also appear within the body of a text, either singly (for example within a footnote), or collected together in a list as a distinct part of a text; detailed bibliographic descriptions of manuscript or other source materials may also be required. These Guidelines propose a number of specialized elements to encode such descriptions, which together constitute the `model.biblLike` class.

`model.biblLike` groups elements containing a bibliographic description.

| | |
|---|---|
| bibl | (bibliographic citation) contains a loosely-structured bibliographic citation of which the sub-components may or may not be explicitly tagged. |
| biblFull | (fully-structured bibliographic citation) contains a fully-structured bibliographic citation, in which all components of the TEI file description are present. |
| biblStruct | (structured bibliographic citation) contains a structured bibliographic citation, in which only bibliographic sub-elements appear and in a specified order. |
| listBibl | (citation list) contains a list of bibliographic citations of any kind. |
| msDesc | (manuscript description) contains a description of a single identifiable manuscript or other text-bearing object such as an early printed book. |

Lists of such elements may also be encoded using the following element:
`<listBibl>` (citation list) contains a list of bibliographic citations of any kind.

In printed texts, the individual constituents of a bibliographic reference are conventionally marked off from each other and from the flow of text by such features as bracketing, italics, special punctuation conventions, underlining, etc. In electronic texts, such distinctions are also important, whether in order to produce acceptably formatted output or to facilitate intelligent retrieval processing,[53] quite apart from the need to distinguish the reference itself as a textual object with particular linguistic properties.

It should be emphasized that for references as for other textual features, the primary or sole consideration is not how the text should be formatted when it is printed or displayed. The distinctions permitted by the scheme outlined here may not necessarily be all that particular formatters or bibliographic styles require, although they should prove adequate to the needs of many such commonly used software systems.[54] The features distinguished and described below (in section *3.12.2. Components of Bibliographic References*) constitute a set which has been useful for a wide range of bibliographic purposes and in many applications, and which moreover corresponds to a great extent with existing bibliographic and library cataloguing practice. For a fuller account of that practice as applied to electronic texts see section *2.2.7. The Source Description*; for a brief mention of related library standards see section *2.8. Note for Library Cataloguers*.

The most commonly used elements in the `model.biblLike` class are <biblStruct> and <bibl>. <biblStruct> will usually be easier to process mechanically than <bibl> because its structure is more constrained and predictable. It is suited to situations in which the objective is to represent bibliographic information for machine processing directly by other systems or after conversion to some other bibliographic markup formats such as BibTeXML or MODS. Punctuation delimiting the components of a print citation is not permitted directly within a <biblStruct> element; instead, the presence and order of child elements must be used to reconstruct the punctuation required by a particular style.

By contrast, <bibl> allows for considerable flexibility in that it can include both delimiting punctuation and unmarked-up text; and its constituents can also be ordered in any way. This makes it suitable for marking up bibliographies in existing documents, where it is considered important to preserve the form of references in the original document, while also distinguishing important pieces of information such as authors, dates, publishers, and so on. <bibl> may also be useful when encoding 'born digital' documents which require use of a specific style guide when rendering the content; its flexibility makes it easier to provide all the information for a reference in the exact sequence required by the target rendering, including any necessary punctuation and linking words, rather than using an XSLT stylesheet or similar to reorder and punctuate the data.

The third element in the `model.biblLike` class, <biblFull>, has a content model based on the <fileDesc> element of the TEI header. Both are based on the International Standard for Bibliographic Description (ISBD), which forms the basis of several national standards for bibliographic citations. The order of child elements in both <biblFull> and <fileDesc> corresponds to the order of bibliographic description 'areas' in ISBD with two minor exceptions. First, the <extent> element, corresponding to the *physical description area* in ISBD, appears just after the *publication, production, distribution, etc. area* in ISBD, not before it as in TEI. Second, <biblFull> and <fileDesc> use the child

[53] For example, to distinguish *London* as an author's name from *London* as a place of publication or as a component of a title.

[54] Among the bibliographic software systems and subsystems consulted in the design of the <biblStruct> structure were BibTeX, Scribe, and ProCite. The distinctions made by all three may be preserved in <biblStruct> structures, though the nature of their design prevents a simple one-to-one mapping from their data elements to TEI elements. For further information, see section *3.12.4. Relationship to Other Bibliographic Schemes*.

element <publicationStmt> to cover not only the *publication, production, distribution, etc. area* but also the *resource identifier and terms of availability area* associated with that publication. Despite these inconsistencies, users encoding citations and attempting to format them according to a standard that closely adheres to ISBD may find that <biblFull>, used with its child elements and without delimiting punctuation, provides an appropriate granularity of encoding with elements that can easily be rendered for the reader. However, it is important to note that some ISBD-derived citation formats (such as ANSI/NISO Z39.29 and ГОСТ 7.1) are not entirely conformant to ISBD either, since they may begin with a statement of authorship that does not map to the ISBD statement of responsibility.

### 3.12.1 Methods of Encoding Bibliographic References and Lists of References

The members of the `model.biblLike` class all share a number of possible component sub-elements. For the <bibl> and <biblStruct> elements, exactly the same sub-elements are concerned, and they are described together in section *3.12.2. Components of Bibliographic References*; for the <biblFull> element, the sub-elements concerned are fully described in section *2.2. The File Description*.

Different levels of specific tagging may be appropriate in different situations. In some cases, it may be felt necessary to mark just the extent of the reference itself, with perhaps a few distinctions being made within it (for example, between the part of the reference which identifies a title or author and the rest). Such references, containing a mixture of text with specialized bibliographic elements, are regarded as <bibl> elements, and tagged accordingly. For example:

```
<p>A book which had a great influence on him
was <bibl>Tufte's <title>Envisioning
       Information</title>
  </bibl>, although he may
never have actually read it.</p>
```

Indeed, some encoders may find it unnecessary to mark the bibliographic reference at all:

```
<p>A book which had a great influence on him
was Tufte's <title>Envisioning Information</title>,
although he may never have actually read it.</p>
```

Some bibliographic references are extremely elliptical, often only a string of the form *Baxter, 1983*. If no further details of Baxter's book are given in the source text and none is supplied by the encoder, then the reference thus given should be tagged as a <bibl>:

```
All of this is of course much more fully treated
in <bibl>Baxter, 1983</bibl>.
```

In general, however, normal modern bibliographic practice, and these Guidelines, distinguish between a bibliographic *reference*, which is a self-sufficient description of a bibliographic item, and a bibliographic *pointer*, which is a short-form citation (e.g. *Baxter, 1983*) which serves usually as a place-holder or pointer to a full long-form reference found elsewhere in the text. The usual encoding of short-form references such as *Baxter, 1983* is not as <bibl> elements but as cross-references to such elements; see section *3.12.3. Bibliographic Pointers* below.

Bibliographic references encoded with <bibl>, <biblFull>, or <biblStruct> may link to another, possibly more detailed, bibliographic reference to the same resource using either the attribute *ref* or *key* defined in `att.canonical`. The *ref* attribute performs this linking function using a URI that may point either within or outside the TEI document, as shown in the following example:

```
<bibl  ref="https://search.worldcat.org/title/song-of-achilles/oclc/911117343">
  <title>The song of Achilles</title> by <author>Madeline Miller</author>.
</bibl>
```

The *key* attribute, on the other hand, associates a bibliographical reference with an external or project-defined identifier as demonstrated here:

```
<bibl key="Homer_Iliad.16">Book 16 of <author>Homer</author>'s <title>Iliad</title>.</bibl>
```

These attributes should not be used for encoding bibliographic pointers (short-form citations); for that purpose, see section *3.12.3. Bibliographic Pointers* below.

In cases where the encoder wishes to impose more structure on the bibliographic information, for example to make sure it conforms to a particular stylesheet or retrieval processor, the <biblStruct> element should be used. Note that several of the features in this and later examples are explained later in the current section.

```
<biblStruct>
  <monogr>
    <author>
      <persName>
        <forename>Edward</forename>
        <forename full="init">R.</forename>
        <surname>Tufte</surname>
      </persName>
      <idno type="scopus">6506403994</idno>
      <idno type="lcaf">https://id.loc.gov/authorities/names/n50012763.html</idno>
    </author>
    <title level="m">Envisioning Information</title>
    <imprint>
      <pubPlace>Cheshire, Conn.</pubPlace>
      <publisher>Graphics Press</publisher>
      <date when="1990"/>
    </imprint>
  </monogr>
</biblStruct>
```

A more complex and detailed bibliographic structure is provided by the <biblFull> element defined in the TEI header module. This element is provided as a means of embedding the file description of one existing digital text within that of another (see further section *2.2. The File Description*); however, its use is not confined to digital texts, and it may be used in the same way as any other bibliographic element, as in this example:

```
<biblFull>
  <titleStmt>
    <title>Envisioning Information</title>
    <author>Tufte, Edward R[olf]</author>
  </titleStmt>
  <extent>126 pp.</extent>
  <publicationStmt>
    <publisher>Graphics Press</publisher>
    <pubPlace>Cheshire, Conn. USA</pubPlace>
    <date>1990</date>
  </publicationStmt>
</biblFull>
```

A list of bibliographic items, of whatever kind, may be treated in the same way as any other list (see section *3.8. Lists*). Alternatively, the specialized <listBibl> element may be used. The difference between the two is that a <list> contains <item> elements, within which bibliographic elements (<bibl>, <biblStruct>, or <biblFull>) may appear, as well as other phrase- and paragraph-level elements, whereas the <listBibl> may contain only bibliographic elements,

optionally preceded by a heading and a series of introductory paragraphs. For most purposes, good practice would usually require that a <listBibl> contain only one kind of bibliographic element, though the following example combines both fully structured <biblStruct> and informal <bibl> elements:

```
<listBibl>
   <head>Bibliography</head>
   <biblStruct  xml:id="NELSON80">
      <analytic>
        <author>
          <persName>
            <surname>Nelson</surname>
            <forename>Theodore</forename>
            <forename>Holm</forename>
          </persName>
        </author>
        <title  level="a">Replacing the printed word:
             a complete literary system</title>
      </analytic>
      <monogr>
         <title  level="m">Information Processing '80:  Proceedings of the IFIPS
             Congress, October 1980</title>
         <editor>
           <persName>
             <forename>Simon</forename>
             <forename>H.</forename>
             <surname>Lavington</surname>
           </persName>
         </editor>
         <imprint>
            <publisher>North-Holland</publisher>
            <pubPlace>Amsterdam</pubPlace>
            <date  when="1980"/>
         </imprint>
         <biblScope  unit="page"  from="1013"
  to="1023">1013–23</biblScope>
      </monogr>
      <note>Apparently a draft of section 4 of
          <title  level="m">Literary Machines</title>.</note>
   </biblStruct>

  <bibl  xml:id="NELSON88"><author><persName><forename>Ted</forename>
<surname>Nelson</surname></persName></author>:
<title  level="u">Literary Machines</title> (privately published,
<date  when="1987">1987</date>).</bibl>

  <bibl  xml:id="BAXTER88"><author><persName><surname>Baxter</surname>,
<forename>Glen</forename></persName></author>:
<title  level="m">Glen Baxter His Life: the years of struggle</title>
<pubPlace>London</pubPlace>: <publisher>Thames and Hudson</publisher>,
<date  when="1988">1988</date>.</bibl>
</listBibl>
```

This example also demonstrates the way that bibliographical markup of authors, titles, dates etc. can be handled differently in <biblStruct>s and <bibl>s. In the two <bibl> items, the key information is marked up, but it is presented in an order which makes it suitable for direct rendering, with the punctuation included.

The <listBibl> element is most appropriate for a more formal bibliography. The same <bibl> or <biblStruct> elements may however be embedded within an ordinary list, thus allowing them to be mixed with running prose or presented informally, as in the following version of the same example:

```
<list>
  <head>Bibliography</head>
  <item>
    <bibl  xml:id="NEL80">
      <author>Nelson, T. H.</author>
      <title  level="a">Replacing the printed word:
            a complete literary system</title>.
    <title  level="m">Information Processing '80:
            Proceedings of the IFIPS Congress, October 1980</title>.
    <editor>Simon H. Lavington</editor>
      <publisher>North-Holland</publisher>:
    <pubPlace>Amsterdam</pubPlace>,
    <date>1980</date>.
    <biblScope>pp 1013–23
      </biblScope>
      <note>Apparently a draft of section 4 of
      <title>Literary Machines</title>.</note>
    </bibl>
  </item>
  <item>
    <bibl  xml:id="NEL88">Ted Nelson: <title>Literary Machines</title>
        (privately published, 1987)</bibl>
  </item>
  <item>
    <bibl  xml:id="BAX88">
      <author>Baxter, Glen</author>
      <title>Glen Baxter His Life: the years of struggle</title>
        London: Thames and Hudson, 1988.
    </bibl>
  </item>
</list>
```

## 3.12.2 Components of Bibliographic References

This section discusses commonly occurring components of bibliographic references and elements used for encoding them. They fall into four groups:

- elements for grouping components of the *analytic*, *monographic*, and *series* levels in a structured bibliographic reference
- titles of various kinds, and statements of intellectual responsibility (authorship, etc.)
- information relating to the publication, pagination, etc. of an item (most of these constitute the default members of the `model.biblPart` class)
- annotation, commentary, and further detail

The following sections describe the elements which may be used to represent such information within a <bibl> or <biblStruct> element. Within the former, elements from the `model.biblPart` class, other phrase-level elements, and plain text may be combined without other constraint; within the latter, such of these elements as exist for a given reference must be distinguished, and must also be presented in a specific order, discussed further below (section *3.12.2.9. Order of Components within References*).

### 3.12.2.1 Analytic, Monographic, and Series Levels

In common library practice a clear distinction is made between an individual item within a larger collection and a free-standing book, journal, or collection. Similarly a book in a series is distinguished sharply from the

series within which it appears. An article forming part of a collection which itself appears in a series thus has a bibliographic description with three quite distinct levels of information:

1. the *analytic* level, giving the title, author, etc., of the article;
2. the *monographic* level, giving the title, editor, etc., of the collection;
3. the *series* level, giving the title of the series, possibly the names of its editors, etc., and the number of the volume within that series.

In the same way, an article in a journal requires at least two levels of information: the analytic level describing the article itself, and the monographic level describing the journal.

A different identifying number may be supplied for any of these three items, that is, for the analytic item, the monographic item, or the series.

Within <bibl>, these three levels may be distinguished simply by the use of the *level* attribute on <title>. They may also be distinguished through the practice of employing nested <bibl> elements. In this example, for instance, the monograph-level component of the reference is encapsulated in its own <bibl> within the main <bibl> for the article:

```
<bibl type="article"
  subtype="magazine_article" xml:id="beaupaire_1911">
  <author>
    <name>
      <surname>Beaupaire</surname>
         (<forename>Edmond</forename>)</name>
  </author>,
<title level="a">A propos de la rue de la Femme-sans-Tête</title>,
<bibl type="monogr">
   <title level="j">La Cité</title>,
  <date when="1911-01">janvier 1911</date>, pp. <biblScope unit="page" from="5" to="17">5-17</b
  </bibl>
</bibl>
```

Within <biblStruct>, the levels are distinguished by the use of the following distinct elements:

`<analytic>` (analytic level) contains bibliographic elements describing an item (e.g. an article or poem) published within a monograph or journal and not as an independent publication.

`<monogr>` (monographic level) contains bibliographic elements describing an item (e.g. a book or journal) published as an independent item (i.e. as a separate physical object).

`<series>` (series information) contains information about the series in which a book or other bibliographic item has appeared.

For purposes of TEI encoding, journals and anthologies are both treated as monographs; a journal title should thus be tagged as a <title level="j"> element within a <monogr> element. Individual articles in the journal or collected texts should be treated at the 'analytic' level. When an article has been printed in more than one journal or collection, the bibliographic reference may have more than one <monogr> element, each possibly followed by one or more <series> elements. A <series> element always relates to the most recently preceding <monogr> element. (Whether reprints of an article are treated in the same bibliographic reference or a separate one varies among different styles. Library lists typically use a different entry for each publication, while academic footnoting practice typically treats all publications of the same article in a single entry.)

The <biblScope> element is used to supply further information about the location of some part of a bibliographic reference. It specifies where to find the component in which it appears within the immediately preceding component of a different level.

In the following example, Schacter's article *Iolaos* appeared on pages 64 to 70 of a volume entitled *Herakles to Poseidon*, which was itself the second of a four volumes published together under the title *Cults of Boitia*; this last title constituted the 38th volume in the series of *Bulletin of the Institute of Classical Studies Supplements*:

```
<biblStruct>
  <analytic>
    <author>Albert Schachter</author>
    <title  level="a">Iolaos</title>
  </analytic>
  <monogr>
    <title  level="m">Herakles to Poseidon</title>
    <imprint>
      <date>1986</date>
    </imprint>
    <biblScope  unit="page">64-70</biblScope>
  </monogr>
  <monogr>
    <title  level="m">Cults of Boiotia</title>
    <imprint>
      <pubPlace>London</pubPlace>
    </imprint>
    <extent>4 vols.</extent>
    <biblScope  unit="part">2</biblScope>
  </monogr>
  <series>
    <title  level="s">Bulletin of the Institute of Classical Studies
        Supplements</title>
    <biblScope  unit="volume">38</biblScope>
  </series>
</biblStruct>
```

In the following example, the article cited has been published twice, once in a journal (where it appeared in volume 40, on pages 3 -46 of the issue of October 1986) and once as a free-standing item, which appeared as number 11 of a German language series.

```
<biblStruct>
  <analytic>
    <author>
      <persName>
        <surname>Thaller</surname>
        <forename>Manfred</forename>
      </persName>
    </author>
    <title  level="a">A Draft Proposal for a Standard for the
        Coding of Machine Readable Sources</title>
  </analytic>
  <monogr>
    <title  level="j">Historical Social Research</title>
    <imprint>
      <date  when="1986-10">October 1986</date>
    </imprint>
    <biblScope  unit="volume">40</biblScope>
    <biblScope  unit="page"  from="3"  to="46">3-46</biblScope>
  </monogr>
  <monogr>
    <title  level="m">Modelling Historical Data:
        Towards a Standard for Encoding and
        Exchanging Machine-Readable Texts</title>
    <editor>
      <persName>
```

```
      <forename>Daniel</forename>
      <forename full="init">I.</forename>
      <surname>Greenstein</surname>
    </persName>
  </editor>
  <imprint xml:lang="de">
    <pubPlace>St. Katharinen</pubPlace>
    <publisher>Max-Planck-Institut für Geschichte
          In Kommission bei
          Scripta Mercaturae Verlag</publisher>
    <date when="1991"/>
  </imprint>
 </monogr>
 <series xml:lang="de">
  <title level="s">Halbgraue Reihe
      zur Historischen Fachinformatik</title>
  <respStmt>
    <resp>Herausgegeben von</resp>
    <name type="person">Manfred Thaller</name>
    <name type="org">Max-Planck-Institut für Geschichte</name>
  </respStmt>
  <title level="s">Serie A: Historische Quellenkunden</title>
  <biblScope unit="volume">11</biblScope>
 </series>
</biblStruct>
```

The practice of analytic vs. monographic citation, as described here, should be distinguished from the practice of including within one citation a reference to another work, which the encoder considers to be related to in some way: see further *3.12.2.7. Related Items* below.

If an identifier is available for the analytic item, it should be represented by means of an <idno> element placed within the <analytic> element, as in the following example where a DOI (Digital Object identifier) is supplied for the article in question.

```
<biblStruct>
 <analytic>
  <author>
    <forename>James</forename>
    <forename>H.</forename>
    <surname>Coombs</surname>
  </author>
  <author>
    <forename>Allen</forename>
    <surname>Renear</surname>
  </author>
  <author>
    <forename>Steven</forename>
    <forename>J.</forename>
    <surname>DeRose</surname>
  </author>
  <title level="a">Markup Systems and The Future of Scholarly Text
      Processing</title>
  <idno type="DOI">10.1145/32206.32209</idno>
  <ref target="http://xml.coverpages.org/coombs.html">http://xml.coverpages.org/coombs.html</re
 </analytic>
 <monogr>
```

```
    <title  level="j">Communications of the ACM</title>
    <imprint>
      <date>1987</date>
    </imprint>
    <biblScope  unit="volume">30</biblScope>
    <biblScope  unit="issue">11</biblScope>
    <biblScope  unit="page">933–947</biblScope>
  </monogr>
</biblStruct>
```

Punctuation must not appear between the elements within a structured bibliographic entry encoded with <biblStruct> or <biblFull>, unless it is contained within the elements it delimits. When (as in most of the examples in this chapter) entries are encoded without any inter-element punctuation, they can be usually be processed more easily by rendering systems able to output bibliographic references in any of several styles.

Within a <bibl> however, it is possible and often convenient to include punctuation.

```
<bibl  xml:id="NELSON_80">
  <author>
    <persName>
      <surname>Nelson</surname>,
    <forename>T.</forename>
      <forename>H.</forename>
    </persName>
  </author>
  <date  when="1980">1980</date>.
<title  level="a">Replacing the printed word: a complete literary
    system</title>. In <title  level="m">Information Processing '80: Proceedings of
the
    IFIPS Congress, October 1980</title>,
ed.
<editor>
    <persName>
      <forename>Simon</forename>
      <forename>H.</forename>
      <surname>Lavington</surname>
    </persName>
  </editor>,
<biblScope  unit="page">1013-23</biblScope>.
<pubPlace>Amsterdam</pubPlace>: <publisher>North-
    Holland</publisher>. (<note>Apparently a draft of section 4 of
  <ref  target="#NELSON_88">
      <title  level="m">Literary
            Machines</title>
    </ref>.</note>)

</bibl>
```

This example shows the components sequenced and punctuated according to the Chicago style, with all the relevant data items marked up appropriately. This markup approach can provide easy rendering, if only one styleguide is targeted, or an original source document uses a specific styleguide, while still allowing for automated recovery of key data items such as names of authors, titles etc.

### 3.12.2.2 Titles, Authors, and Editors

Bibliographic references typically include the title of the work being cited and the names of those intellectually responsible for it. For articles in journals or collections, such statements should appear both for the analytic and for the monographic level. The following elements are provided for tagging such elements:

`<title>` (title) contains a title for any kind of work.

`<author>` (author) in a bibliographic reference, contains the name(s) of an author, personal or corporate, of a work; for example in the same form as that provided by a recognized bibliographic name authority.

`<editor>` contains a secondary statement of responsibility for a bibliographic item, for example the name of an individual, institution or organization, (or of several such) acting as editor, compiler, translator, etc.

`<respStmt>` (statement of responsibility) supplies a statement of responsibility for the intellectual content of a text, edition, recording, or series, where the specialized elements for authors, editors, etc. do not suffice or do not apply. May also be used to encode information about individuals or organizations which have played a role in the production or distribution of a bibliographic work.

`<resp>` (responsibility) contains a phrase describing the nature of a person's intellectual responsibility, or an organization's role in the production or distribution of a work.

`<name>` (name, proper noun) contains a proper noun or noun phrase.

`<meeting>` contains the formalized descriptive title for a meeting or conference, for use in a bibliographic description for an item derived from such a meeting, or as a heading or preamble to publications emanating from it.

`<sponsor>` (sponsor) specifies the name of a sponsoring organization or institution.

`<funder>` (funding body) specifies the name of an individual, institution, or organization responsible for the funding of a project or text.

`<distributor>` (distributor) supplies the name of a person or other agency responsible for the distribution of a text.

`<principal>` (principal researcher) supplies the name of the principal researcher responsible for the creation of an electronic text.

The elements <author>, <editor>, <respStmt>, <meeting>, <sponsor>, <funder>, and <principal> are the default members of the `model.respLike` class, a subclass of the `model.biblPart` class to which the constituents of the <bibl> element belong.

In bibliographic references, all titles should be tagged as such, whether analytic, monographic, or series titles. The single element <title> is used for all these cases. When it appears directly within an <analytic>, <monogr>, or <series> element, <title> is interpreted as belonging to the appropriate level. However, it is recommended that the *level* attribute be used to signal this explicitly.

It is a semantic error to give a value for the *level* attribute which is inconsistent with the context. The *level* value a implies the analytic level; the values m, j, and u imply the monographic level; the value s implies the series level. Note, however, that the semantic error occurs only if the nested title is directly enclosed by the <analytic>, <monogr>, or <series> element; if it is enclosed only indirectly (i.e., nested more deeply), no semantic error need be present. For example, the analytic title may contain a monographic title, as in the following example:

```
<biblStruct>
  <analytic>
    <author  ref="http://id.loc.gov/authorities/names/no2001067434">
      <persName>
        <forename>Lucy</forename>
        <forename>Allen</forename>
        <surname>Paton</surname>
      </persName>
    </author>
    <title>Notes on Manuscripts of the
    <title  level="m"  xml:lang="fr">Prophécies de Merlin</title>
    </title>
  </analytic>
  <monogr>
    <title  level="j">PMLA</title>
    <imprint>
      <date>1913</date>
    </imprint>
```

```
    <biblScope  unit="volume">8</biblScope>
    <biblScope  unit="page">122</biblScope>
  </monogr>
</biblStruct>
```

In this case, the analytic title 'Notes on Manuscripts of the *Prophécies de Merlin*' needs no *level* attribute because it is directly contained by an <analytic> element; the monographic title contained within it, 'Prophécies de Merlin', is not semantically erroneous because it is not directly contained by the <analytic> element.

In some bibliographic applications, it may prove useful to distinguish main titles from subordinate titles, parallel titles, etc. The *type* attribute is provided to allow this distinction to be recorded.

The following reference, from a national standard for bibliographic references, illustrates this type of analysis with its distinction between main and subordinate titles. Note that this uses the more flexible <bibl>, rather than the structured <biblStruct> element: consequently, there is no requirement to tag all the components of the reference (notably the authors).

```
<bibl>Saarikoski, Pirkko-Liisa, and Paavo Suomalainen,
<title  level="a"  type="main">Studies on the physiology of
    the hibernating hedgehog, 15</title>
  <title  level="a"  type="sub">Effects of seasonal
    and temperature changes on the in vitro glycerol release from
    brown adipose tissue</title>
  <title  level="j">Ann. Acad. Sci. Fenn., Ser. A4</title>
  <date>1972</date>
  <biblScope  unit="volume">187</biblScope>
  <biblScope  unit="page"  to="4">1-4</biblScope>
</bibl>
```

Slightly more complex is the distinction made below among main, subordinate, and parallel titles, in an example from the same source (p. 63). The punctuation and the bibliographic analysis are those given in ANSI Z39.29-1977; the punctuation is in the style prescribed by the International Standard Bibliographic Description (ISBD).[55] Again, it is only because this example uses <bibl> rather than <biblStruct>, that specific punctuation may be included between the component elements of the reference.

```
<bibl>Tchaikovsky, Peter Ilich.
<title  level="m"  type="main">The swan lake ballet</title>
= <title  level="m"  type="parallel"
   xml:lang="fr">Le lac des cygnes</title>
: <title  level="m"  type="sub"  xml:lang="fr">grand ballet en 4 actes</title>
: <title  level="m"  type="sub">op. 20</title>
[Score].
New York: Broude Brothers; [1951] (B.B. 59). vi, 685 p.</bibl>
```

The elements <author> and <editor> have fairly obvious significance for printed books and articles; for other kinds of bibliographic items their proper usage may be less obvious. The <author> element should be used for the person or agency with primary responsibility for a work's intellectual content, and the element <editor> for other people or agencies with some responsibility for that content, whether or not they are called 'editor'. An organization such as a radio or television station is usually accounted 'author' of a broadcast, for example, while the author of a government report will usually be the agency which produced it. A translator, illustrator, or compiler, may however be marked by means of the <editor> element, optionally using the *role* attribute to specify the nature of their responsibility more exactly.

[55] The analysis is not wholly unproblematic: as the text of the standard points out, the first subordinate title is subordinate only to the parallel title in French, while the second is subordinate to both the English main title and the French parallel title, without this relationship being made clear, either in the markup given in the example or in the reference structure offered by the standard.

Many bibliographic and Linked Data applications require disambiguation of author names using unique identifiers. Both the <author> and <editor> elements may contain one or more <idno> elements, to supply such identifiers. Alternatively, if only a single identifier is to be recorded, the *key* or *ref* attribute may be used, as further discussed in *3.6.1. Referring Strings*.

```
<bibl>
  <author  ref="http://viaf.org/viaf/95301405">John Warrack</author>. „Es waren seine
letzten Töne!“
In <editor  ref="http://viaf.org/viaf/263865979">Joachim Veit</editor>
and <editor  ref="http://viaf.org/viaf/268371810">Frank Ziegler</editor> eds.
Weber-Studien Bd. 3, Mainz (1996), pp.300–317

</bibl>
```

For anyone else with responsibility for the work, the <respStmt> element should be used. The nature of the responsibility is indicated by means of a <resp> element, and the person, organization, etc. responsible by a <name>, <persName>, or <orgName> element. Strings such as ‘unknown’ may be encoded using the <rs> element. A <respStmt> should comprise either at least one of the four naming elements (<name>, <persName>, <orgName>, or <rs>) followed by one or more <resp> elements, or at least one <resp> element followed by one or more of the four naming elements.

Examples of secondary responsibility of this kind include the roles of illustrator, translator, encoder, and annotator. The <respStmt> element may also be used for editors, if it is desired to record the specific terms in which their role is described.

Examples of <author> and <editor> may be found in sections *3.12.1. Methods of Encoding Bibliographic References and Lists of References*, and *3.12.2.1. Analytic, Monographic, and Series Levels*; wherever <author> and <editor> may occur, the <respStmt> element may also occur. When one of these elements precedes or immediately follows a title, it applies to that title; when it follows an <edition> element or occurs within an edition statement, it applies to the edition in question.

In this example, the <respStmt> elements apply to the work as a whole, not merely to the first edition:

```
<bibl>
  <author>Lominandze, DG</author>.
<title  level="m">Cyclotron waves in plasma</title>.
<respStmt>
   <resp>Translated by</resp>
   <name>AN. Dellis</name>
  </respStmt>;
<respStmt>
   <resp>edited by</resp>
   <name>SM. Hamberger</name>
  </respStmt>.
<edition>1st ed.</edition>
  <pubPlace>Oxford</pubPlace>:
<publisher>Pergamon Press</publisher>,
<date>1981</date>.
<extent>206 p.</extent>
  <title  level="s">International series in natural philosophy</title>.
<note  place="inline">Translation of:
  <title  xml:lang="ru-Latn"  level="m">Ciklotronnye volny v
       plazme</title>.
  <idno  type="ISBN">0-08-021680-3</idno>.
  </note>
</bibl>
```

This example retains the original punctuation and editorial conventions of the source (ISO 690:1987) and is therefore encoded using the <bibl> element.

In the following example, by contrast, the <respStmt> element applies to the edition, and not to the collection *per se* (Moser and Tervooren were not responsible for the first thirty-five printings). As is permissible within a <biblStruct> element, the component elements have been reordered from their appearance on the title page of the volume in order to ensure the correct relationship of the collection title, the edition statement, and the statement of responsibility.

```
<biblStruct>
  <monogr  xml:lang="de">
    <title  level="m">Des Minnesangs Frühling</title>
    <note  place="inline">Mit 1 Faksimile</note>
    <edition>36., neugestaltete und erweiterte Auflage</edition>
    <respStmt>
      <resp>Unter Benutzung der Ausgaben von <name>Karl
                Lachmann</name> und <name>Moriz Haupt</name>, <name>Friedrich
                Vogt</name> und <name>Carl von Kraus</name> bearbeitet von</resp>
      <name>Hugo Moser</name>
      <name>Helmut Tervooren</name>
    </respStmt>
    <imprint>
      <pubPlace>Stuttgart</pubPlace>
      <publisher>S. Hirzel Verlag</publisher>
      <date>1977</date>
    </imprint>
    <biblScope  unit="volume">I Texte</biblScope>
  </monogr>
</biblStruct>
```

The party with a particular responsibility for the intellectual content may vary over time. Likewise, a given individal's responsibility or role may change over time. These situations may be recorded with the <respStmt> element. For example, the following could be used when one proofreader took over for another.

```
<respStmt>
  <resp>proofreading</resp>
  <persName  from="1994-02"  to="1994-05">Ashley Cross</persName>
  <persName  from="1994-06"  to="1994-10">Loren Noveck</persName>
</respStmt>
```

The following example records the fact that one individual had two distinctly different intellectual responsibilities at different times.

```
<respStmt>
  <persName>Erica Dillon</persName>
  <resp  when="2000-08">annotated uncredited citations</resp>
  <resp  when="2001-03">encoded named entities</resp>
</respStmt>
```

Another form of 'responsibility' arises when a work is published as the outcome of a conference, workshop or similar meeting. The <meeting> element may be used to supply this information, as in the following example:

```
<biblStruct>
  <monogr>
    <title  level="m">Proceedings of a workshop on corpus resources</title>
```

```
    <respStmt>
      <resp>Programme Organizer</resp>
      <name>Geoffrey Leech</name>
    </respStmt>
    <meeting>DTI Speech and Language Technology Club meeting, 3-4
        January 1990, Wadham College, Oxford</meeting>
    <imprint>
      <pubPlace>Oxford</pubPlace>
    </imprint>
  </monogr>
</biblStruct>
```

#### 3.12.2.3 Document Identifiers

Many bibliographic references include identifiers for a work to help with precise identification of an appropriate document. For example, a book in the *Short Title Catalogue* could be referenced with its STC number:

```
<biblStruct>
  <monogr>
    <author>
      <forename>John</forename>
      <surname>Downame</surname>
    </author>
    <title  type="short">Foure treatises tending to disswade all Christians from foure
no lesse hainous then common sinnes</title>
    <idno  type="stc2ndEd">7141</idno>
    <imprint>
      <pubPlace>At London</pubPlace>
      <publisher>Imprinted by Felix Kyngston, for William Welby, and are to be sold at
his shop in Pauls Church-yard at the signe of the Greyhound</publisher>
      <date  when="1609">1609</date>
    </imprint>
  </monogr>
</biblStruct>
```

However, some bibliographic references actually *require* identifiers of various types because they do not include a statement of the title and the names of those intellectually responsible for it. The following elements may be used for such purposes:

`<orgName>` (organization name) contains an organizational name.

`<idno>` (identifier) supplies any form of identifier used to identify some object, such as a bibliographic item, a person, a title, an organization, etc. in a standardized way.

`<classCode>` (classification code) contains the classification code used for this text in some standard classification system.

`<date>` (date) contains a date in any format.

For example, a citation to a patent typically includes a country or organization code (a two-character code identifying a patent authority) and a serial number for the patent (whose structure varies by patent authority). The citation might also contain a *kind code* (which characterizes a particular publication for the patent and which corresponds to a specific stage in the patent procedure) and the date when the patent was filed with or published by the issuing authority. For bibliographic references to patents, the above elements may be used as follows:

- <orgName>, within <authority>, may be used to contain the code of the patent authority. The *type* attribute may be used to specify the type of patent authority (such as a national patent office or a supra-national patent organization).
- <idno> may be used to contain the serial number assigned by the corresponding patent authority.
- <classCode> may be used to contain the kind code of the patent document.

- <date> may be used to contain the date of the patent document. The *type* attribute may be used to specify whether this corresponds to the filing date of a patent application or the publication date of a patent publication.

The following reference illustrates an encoding for a patent publication which might be cited in print as 'United States patent US 6,885,550 B1, issued April 26, 2005':

```
<biblStruct type="patent"
  status="publication">
  <monogr>
    <authority>
      <orgName type="national">US</orgName>
    </authority>
    <idno type="docNumber">6885550</idno>
    <imprint>
      <classCode scheme="http://www.uspto.gov/">B1</classCode>
      <date type="publicationDate"
        when="2005-04-26">April 26, 2005</date>
    </imprint>
  </monogr>
</biblStruct>
```

#### 3.12.2.4 Imprint, Size of a Document, and Reprint Information

By *imprint* is meant all the information relating to the publication of a work: the person or organization by whose authority and in whose name a bibliographic entity such as a book is made public or distributed (whether a commercial publisher or some other organization), the place and the date of publication. It may also include a full address for the publisher or organization. A full bibliographic references will usually also specify the number of pages in a print publication (or equivalent information for non-print materials), and possibly also the specific location of the material being cited within its containing publication. The following elements are provided to hold this information:

`<imprint>` groups information relating to the publication or distribution of a bibliographic item.

`<address>` (address) contains a postal address, for example of a publisher, an organization, or an individual.

`<pubPlace>` (publication place) contains the name of the place where a bibliographic item was published.

`<publisher>` (publisher) provides the name of the organization responsible for the publication or distribution of a bibliographic item.

`<date>` (date) contains a date in any format.

`<extent>` (extent) describes the approximate size of a text stored on some carrier medium or of some other object, digital or non-digital, specified in any convenient units.

`<idno>` (identifier) supplies any form of identifier used to identify some object, such as a bibliographic item, a person, a title, an organization, etc. in a standardized way.

Members of the model classes `model.imprintPart` and `model.dateLike` may appear inside an <imprint> element in a specific location within a <biblStruct>, or alternatively, they may appear alongside any other bibliographic component inside a <bibl>.

`model.imprintPart` groups the bibliographic elements which occur inside imprints.

| | |
|---|---|
| biblScope | (scope of bibliographic reference) defines the scope of a bibliographic reference, for example as a list of page numbers, or a named subdivision of a larger work. |
| distributor | (distributor) supplies the name of a person or other agency responsible for the distribution of a text. |
| publisher | (publisher) provides the name of the organization responsible for the publication or distribution of a bibliographic item. |
| pubPlace | (publication place) contains the name of the place where a bibliographic item was published. |

`model.dateLike` groups elements containing temporal expressions.

| | |
|---|---|
| date | (date) contains a date in any format. |
| time | (time) contains a phrase defining a time of day in any format. |

For bibliographic purposes, usually only the place (or places) of publication are required, possibly including the name of the country, rather than a full address; the element <pubPlace> is provided for this purpose. Where however the full postal address is likely to be of importance in identifying or locating the bibliographic item concerned, it may be supplied and tagged using the <address> element described in section *3.6.2. Addresses.* Alternatively, if desired, the <rs> or <name> elements described in section *3.6.1. Referring Strings* may be used; this involves no claim that the information given is either a full address or the name of a city.

The name of the publisher of an item should be marked using the <publisher> element even if the item is made public ('published') by an organization other than a conventional publisher, as is frequently the case with technical reports:

```
<biblStruct>
  <monogr>
    <author>Nicholas, Charles K.</author>
    <author>Welsch, Lawrence A.</author>
    <title  level="m">On the interchangeability of SGML and ODA</title>
    <idno  type="NIST">NISTIR 4681</idno>
    <imprint>
      <pubPlace>Gaithersburg, MD</pubPlace>
      <publisher>National Institute of Standards and Technology
      </publisher>
      <date  when="1992-01">January 1992</date>
    </imprint>
    <extent>19 pp.</extent>
  </monogr>
</biblStruct>
```

and with dissertations:

```
<biblStruct>
  <monogr>
    <author>Hansen, W.</author>
    <title  level="u">Creation of hierarchic text
        with a computer display</title>
    <idno  type="ANL">ANL-7818</idno>
    <note  place="inline">Ph.D. dissertation</note>
    <imprint>
      <publisher>Dept. of Computer Science, Stanford Univ.</publisher>
      <pubPlace>Stanford, CA</pubPlace>
      <date  when="1971-06">June 1971</date>
    </imprint>
  </monogr>
</biblStruct>
```

In this second example, the <idno> element is used to provide the identifier allocated to the thesis by the Argonne National Laboratory. Since it applies to the monographic element, the <idno> should be provided as a direct child of the <monogr> element, rather than elsewhere in the <biblStruct> element.

The specialist elements <publisher> and <distributor> are provided to cover the most common roles related to the production and distribution of a bibliographical item, but other roles such as printer and bookseller may also need to be encoded, and <respStmt> is available inside <imprint> for this purpose.

When an item has been reprinted, especially reprinted without change from a specific earlier edition, the reprint may appear in a <monogr> element with only the <imprint> and other details of the reprint. In the following

example, a microform reprint has been issued without any change in the title or authorship. The series statement here applies only to the second <monogr> element.

```
<biblStruct>
  <monogr>
    <author>Shirley, James</author>
    <title type="main">The gentlemen of Venice</title>
    <title type="sub">a tragi-comedie presented at the private
        house in Salisbury Court by Her Majesties servants</title>
    <note place="inline">[Microform]</note>
    <imprint>
      <pubPlace>London</pubPlace>
      <publisher>H. Moseley</publisher>
      <date>1655</date>
    </imprint>
    <extent>78 p.</extent>
  </monogr>
  <monogr>
    <imprint>
      <pubPlace>New York</pubPlace>
      <publisher>Readex Microprint</publisher>
      <date>1953</date>
    </imprint>
    <extent>1 microprint card, 23 x 15 cm.</extent>
  </monogr>
  <series>
    <title level="s">Three centuries of drama: English, 1642–1700</title>
  </series>
</biblStruct>
```

This encoding can be extended to the case of patent documents, where the same patent application is published, with or without changes, at different stages of the patenting procedure. In this case, the kind code and, optionally, the publication date characterize different publications of the same patent application during the procedure. For example:

```
<biblStruct type="patent"
  status="publication">
  <monogr>
    <authority>
      <orgName type="national">EP</orgName>
    </authority>
    <idno type="docNumber">1558513</idno>
    <imprint>
      <classCode scheme="http://www.epo.org/">A1</classCode>
      <date type="publicationDate"
        when="2005-08-03"/>
    </imprint>
  </monogr>
  <monogr>
    <imprint>
      <classCode scheme="http://www.epo.org/">B1</classCode>
      <date type="publicationDate"
        when="2009-09-09"/>
    </imprint>
  </monogr>
</biblStruct>
```

The above bibliographic reference discloses different publications of the patent EP1558513 during the patenting procedure. The first publication from 3 August 2005 has the kind code ”A1” indicating that it is a published patent application comprising the European search report issued after carrying out the search at the European Patent Office, whereas the second publication from 9 September 2009 has the kind code ”B1” indicating that it was published after the patent application has been granted.

An alternative way of handling the above situations would be to use the <relatedItem> element described in section *3.12.2.7. Related Items* below.

### 3.12.2.5 Scopes and Ranges in Bibliographic Citations

Many bibliographic citations contain data limiting the citation to one or more volumes, issues, or pages, or to a name or number of a subdivison of the host work. These come in two varieties:

- the scope of a bibliographic reference (encoded using <biblScope>)
- the range of a work cited (encoded using <citedRange>)

Where it is desired to distinguish different classes of such information (volume number, page number, chapter number, etc.), the *unit* attribute may be used with any convenient typology (see the element definitions for <biblScope> and <citedRange> for some suggested values).

A scope of a bibliographic reference defines that the *entire work cited* may be found in particular volumes, issues, pages, etc. For example:

```
<biblStruct>
  <analytic>
    <author>
      <persName>
        <surname>Wrigley</surname>
        <forename  full="init">E.</forename>
        <forename  full="init">A.</forename>
      </persName>
    </author>
    <title  level="a">Parish registers and the historian</title>
  </analytic>
  <monogr>
    <author>
      <persName>
        <surname>Steel</surname>
        <forename  full="init">D.</forename>
        <forename  full="init">J.</forename>
      </persName>
    </author>
    <author>
      <persName>
        <surname>Steel</surname>
        <forename  full="init">A.</forename>
        <forename  full="init">E.</forename>
        <forename  full="init">F.</forename>
      </persName>
    </author>
    <title  level="m">General sources of births, marriages and deaths before
1837</title>
    <imprint>
      <pubPlace>London</pubPlace>
      <publisher>Society of Genealogists</publisher>
      <date  when="1968"/>
    </imprint>
    <biblScope  unit="page"  from="155"  to="167">155–167</biblScope>
  </monogr>
  <series>
```

```
    <title  level="s">National index of parish registers</title>
    <biblScope  unit="volume">1</biblScope>
  </series>
</biblStruct>
```

The *unit* attribute on <biblScope> is optional: both the following are legal examples:

```
<biblStruct>
  <analytic>
    <author>Boguraev, Branimir</author>
    <author>Neff, Mary</author>
    <title  level="a">Text Representation, Dictionary Structure,
        and Lexical Knowledge</title>
  </analytic>
  <monogr>
    <title  level="j">Literary & Linguistic Computing</title>
    <imprint>
      <date>1992</date>
    </imprint>
    <biblScope  unit="volume">7</biblScope>
    <biblScope  unit="issue">2</biblScope>
    <biblScope  unit="page">110-112</biblScope>
  </monogr>
</biblStruct>
```

```
<biblStruct>
  <analytic>
    <author>Chesnutt, David</author>
    <title  level="a">Historical Editions in the States</title>
  </analytic>
  <monogr>
    <title  level="j">Computers and the Humanities</title>
    <imprint>
      <date  when="1991-12">(December, 1991):</date>
    </imprint>
    <biblScope>25.6</biblScope>
    <biblScope  from="377"  to="380">377–380</biblScope>
  </monogr>
</biblStruct>
```

On the other hand, a cited range encodes that the author *cited only the portion* defined by this range. For example, a footnote following a quotation from page 378 of 'Historical Editions in the States' that includes a full bibliographic reference would be encoded using <biblStruct> as follows:

```
<biblStruct>
  <analytic>
    <author>Chesnutt, David</author>
    <title  level="a">Historical Editions in the States</title>
  </analytic>
  <monogr>
    <title  level="j">Computers and the Humanities</title>
    <imprint>
```

```
      <date when="1991-12">(December, 1991):</date>
    </imprint>
    <biblScope>25.6</biblScope>
    <biblScope unit="page" from="377" to="380">377–380</biblScope>
  </monogr>
  <citedRange>378</citedRange>
</biblStruct>
```

#### 3.12.2.6 Series Information

Series information may (in <bibl> elements) or must (in <biblStruct> elements) be enclosed in a <series> element or (in a <biblFull> element) a <seriesStmt> element. The title of the series may be tagged <title level="s">, the volume number <biblScope unit="volume">, and responsibility statements for the series (e.g. the name and affiliation of the editor, as in the example in section *3.12.2.1. Analytic, Monographic, and Series Levels*) may be tagged <editor> or <respStmt>. Any identifier associated with the series itself should be marked using the <idno> element.

#### 3.12.2.7 Related Items

In bibliographic parlance, a *related item* is any bibliographic item which, though related to that being defined, is distinct from it. The distinction between analytic and monographic items made above may be thought of as a special case of this kind of 'related' item. More usually however, the term is applied to such items as translations, continuations, different versions, parts, etc.

The element <relatedItem> is provided as a means of documenting such associated items:

`<relatedItem>` contains or references some other bibliographic item which is related to the present one in some specified manner, for example as a constituent or alternative version of it.

In the following example, the first <biblStruct> describes a facsimile edition, and the second describes the work of which it is a facsimile. The relation between the facsimile and its source is represented by means of a <relatedItem> within the first description, which points to the description of the source.

```
<biblStruct xml:id="bibl03">
  <monogr>
    <author>Swinburne, Algernon Charles</author>
    <title level="m">Swinburne's <title level="m">Atalanta in Calydon</title>: A
Facsimile of the
        First Edition</title>
    <editor>Georges Lafourcade</editor>
    <imprint>
      <pubPlace>London</pubPlace>
      <publisher>Oxford UP</publisher>
      <date>1930</date>
    </imprint>
  </monogr>
  <relatedItem type="otherEdition">
    <ref target="#bibl04"/>
  </relatedItem>
</biblStruct>
<biblStruct xml:id="bibl04">
  <monogr>
    <author> Swinburne, Algernon Charles</author>
    <title level="m">Atalanta in Calydon</title>
    <imprint>
      <pubPlace>London</pubPlace>
      <publisher>Edward Moxon</publisher>
      <date>1865</date>
    </imprint>
  </monogr>
</biblStruct>
```

The <ref> element in the above example could be replaced by the referenced <biblStruct> itself since a <relatedItem> may contain any form of bibliographic reference. For example, one of the examples quoted above might also be encoded as follows:

```
<biblStruct>
  <monogr>
    <author>Shirley, James</author>
    <title  type="main">The gentlemen of Venice</title>
    <imprint>
      <pubPlace>New York</pubPlace>
      <publisher>Readex Microprint</publisher>
      <date>1953</date>
    </imprint>
    <extent>1 microprint card, 23 x 15 cm.</extent>
  </monogr>
  <series>
    <title  level="s">Three centuries of drama: English, 1642–1700</title>
  </series>
  <relatedItem  type="otherEdition">
    <biblStruct>
      <monogr>
        <author>Shirley, James</author>
        <title  type="main"  level="m">The gentlemen of Venice</title>
        <title  type="sub"  level="m">a tragi-comedie presented at the private
               house in Salisbury Court by Her Majesties servants</title>
        <imprint>
          <pubPlace>London</pubPlace>
          <publisher>H. Moseley</publisher>
          <date  when="1655">1655</date>
        </imprint>
        <extent>78 p.</extent>
      </monogr>
    </biblStruct>
  </relatedItem>
</biblStruct>
```

The *type* attribute should be used to indicate the relationship between the bibliographic item and any <relatedItem> it contains or points to. The relationships may be transitive (for example translatedAs or reprintedFrom) or non-transitive (for example otherEdition). The *subtype* attribute may be used to provide a more detailed classification, where this is appropriate. Some further examples follow:

```
<biblStruct>
  <monogr>
    <author>Tolkien, J.R.R.</author>
    <title  level="m">Den hobbit</title>
    <title  type="sub">aus dem Engleschen iwwersat</title>
    <editor  role="translator">Henry Wickens</editor>
    <imprint>
      <pubPlace>Esch-sur-Sûre</pubPlace>
      <publisher>Op der Lay S. àr. L</publisher>
      <date>2002</date>
    </imprint>
  </monogr>
  <relatedItem  type="translatedFrom">
    <bibl>
      <author>Tolkien, J.R.R.</author>
```

```
      <title level="m">The Hobbit</title>.
    <publisher>Collins</publisher>
      <date>1997</date>
    </bibl>
  </relatedItem>
</biblStruct>
```

In this example, a full bibliographic description of the edition used as source for the translation is provided within the content of the <relatedItem>. Alternatively this might be provided by means of a link, in which case the <relatedItem> would be empty:

```
<relatedItem type="translatedFrom"
  target="http://www.example.com/bibliography.xml#TOLK97"/>
```

#### 3.12.2.8 Notes and Statement of Language

Explanatory notes about the publication of unusual items, the form of an item (e.g. *[Score]* or *[Microform]*), or its provenance (e.g. *translation of …*) may be tagged using the <note> element. The same element may be used for any descriptive annotation of a bibliographic entry in a database.

`<note>` (note) contains a note or annotation.

For example:

```
<bibl>
  <author>Coombs, James H., Allen H. Renear,
    and Steven J. DeRose.</author>
  <title level="a">Markup Systems and the Future of Scholarly
    Text Processing.</title>
  <title level="j">Communications of the ACM</title>
  <biblScope>30.11 (November 1987): 933–947.</biblScope>
  <note>Classic polemic supporting descriptive over procedural
    markup in scholarly work.</note>
</bibl>
```

The <textLang> element may be used to record information about the languages used within a bibliographic item.

`<textLang>` (text language) describes the languages and writing systems identified within the bibliographic work being described, rather than its description.

- `@mainLang` (main language) supplies a code which identifies the chief language used in the bibliographic work.
- `@otherLangs` (other languages) one or more codes identifying any other languages used in the bibliographic work.

This element can take the form of a simple note such as:

```
<textLang>Latin, with some glosses in Anglo-Saxon and French</textLang>
```

However, it is generally recommended where feasible to use the *mainLang* attribute to record the chief language of the bibliographic item, and optionally the *otherLangs* to identify other languages used in the work. For example:

```
<textLang mainLang="la" otherLangs="ang fr">Latin, with some glosses in Anglo-Saxon
and French</textLang>
```

The *mainLang* and *otherLangs* attributes should both provide language identifiers in the same form as used for *xml:lang* as described at *vi.1 Language Identification*. Where additional detail is needed correctly to describe a language, or to discuss its deployment in a given text, this should be done using the <langUsage> element in the TEI header, within which individual <language> elements document the languages used: see *2.4.2. Language Usage*.

A description, in French, of a work predominantly in German, but also with some Latin might have a <textLang> like the following:

```
<textLang  xml:lang="fr"  mainLang="de"
  otherLangs="la">allemand et latin</textLang>
```

For more information about the use of <textLang> in manuscript descriptions see: *11.6.6. Languages and Writing Systems*.

#### 3.12.2.9 Order of Components within References

The order of elements in <bibl> elements is not constrained.

In <biblStruct> elements, the <analytic> element, if it occurs, must come first, followed by one or more <monogr> and <series> elements, which may appear intermingled (as long as a <monogr> element comes first), and then zero or more of the following in any order: <note>, <witDetail>, <idno>, <ptr>, <ref>, <relatedItem>, and <citedRange>. Within <analytic>, the title(s), author(s), editor(s), and other statements of responsibility may appear in any order; it is recommended that all forms of the title be given together. Within <monogr>, the author, editor, and statements of responsibility may either come first or else follow the monographic title(s). Following these, the elements listed below, if present, must appear in the following order:

- <note>s on the publication (and <meeting> elements describing the conference, in the case of a proceedings volume)
- <edition> elements, each followed by any related <editor> or <respStmt> elements
- <imprint>
- <biblScope>

Within <imprint>, the elements allowed may appear in any order.

Finally, within the <series> information in a <biblStruct>, the sequence of elements is not constrained.

If more detailed structuring of a bibliographic description is required, the <biblFull> element should be used. This is not further described here, as its contents are essentially equivalent to those of the <fileDesc> element in the <teiHeader>, which is fully described in section *2.2. The File Description*.

### 3.12.3 Bibliographic Pointers

References which are pointers to bibliographic items, of whatever kind, should be treated in the same way as other cross-references (see section *3.7. Simple Links and Cross-References*). As discussed in that section, cross-referencing within TEI texts is in general represented by means of <ptr> or <ref> elements. A *target* attribute on these elements is used to supply an identifying value for the target of the cross-reference, which should be, in the case of bibliographic elements, a bibliographic reference of some kind. Where the form of the reference itself is unimportant, or may be reconstructed mechanically, or is not to be encoded, the <ptr> element is used, as in the following example:

```
As shown above (<ptr  target="#NEL80"/>) ...
```

Where the form of the reference is important, or contains additional qualifying information which is to be kept but distinguished from the surrounding text, the <ref> element should be used, as in the following example:

```
Nelson claims <ref  target="#NEL80">(ibid, passim)</ref> ...
```

It may be important to distinguish between the short form of a bibliographic reference and some qualifying or additional information. The latter should not appear within the scope of the <ref> element when this is the case, as for example in an application concerned to normalize bibliographic references:

```
Nelson claims (<ref  target="#NEL80">Nelson [1980]</ref> pages 13–37)
...
```

If it is desired to capture additional information like this in a short-form reference, then <bibl> may be used with the *ref* attribute pointing to the full bibliographic reference:

```
Nelson claims (<bibl  ref="#NEL80">Nelson [1980] pages <biblScope  unit="page"  from="13"  to="37">
37</biblScope>
</bibl>) ...
```

The <ref> element may also be used to provide a reference to a copy of the bibliographic item itself, particularly if this is available online, as in the following example:

```
<biblStruct>
  <analytic>
    <author>
      <forename>Suzana</forename>
      <surname>Sukovic</surname>
    </author>
    <title  level="a">Beyond the Scriptorium: The Role of the Library in Text
        Encoding</title>
    <ref  target="https://www.dlib.org/dlib/january02/sukovic/01sukovic.html">https://www.dlib.org
  </analytic>
  <monogr>
    <title  level="j">D-Lib</title>
    <imprint>
      <biblScope  unit="volume">8</biblScope>
      <biblScope  unit="issue">1</biblScope>
      <date>2002</date>
    </imprint>
  </monogr>
</biblStruct>
```

The <ptr> element may be used as a child element of <biblStruct> to refer to the online catalog record of this bibliographic item:

```
<biblStruct  type="manifestation">
  <monogr>
    <author>
      <persName>
        <forename>Germain</forename>
        <surname>Brice</surname>
      </persName>
    </author>
    <title  level="m">Description de la ville de Paris et de tout ce qu’elle contient
de plus remarquable, par Germain Brice ; enrichie d’un nouveau plan et de figures
dessinées et gravées correctement. 7e édition, revue et augmentée par l’auteur</title>
    <imprint>
      <date  when="1717">1717</date>
      <pubPlace>Paris</pubPlace>
      <publisher>F. Fournier</publisher>
    </imprint>
    <extent>In-12</extent>
```

```
  </monogr>
  <ptr  type="catBnf"
    target="http://catalogue.bnf.fr/ark:/12148/cb30160624f/"/>
</biblStruct>
```

### 3.12.4 Relationship to Other Bibliographic Schemes

The bibliographic tagging defined here can capture the distinctions required by most bibliographic encoding systems; for the benefit of users of some commonly used systems, the following lists of equivalences are offered, showing the relationship of the markup defined here to the fields defined for bibliographic records in the Scribe, BibTeX, and ProCite systems.

Listed below are the equivalences between the various bibliographic fields defined for use in the Scribe and BibTeX systems of bibliographic databases and the elements defined in this module.[56] Elements and structures available in the module defined here which have no analogues in Scribe and BibTeX are not noted.

**address** tag as <placeName> or <address>

**annote** tag as <note>

**author** tag as <author>

**booktitle** tag as <title level="m"> or <title> within <monogr>

**chapter** tag as <biblScope unit="chap">

**date** used only to record date entry was made in the bibliographic database; not supported

**edition** tag as <edition>

**editor** tag as <editor> or <respStmt>

**editors** tag as multiple <editor> or <respStmt> elements

**fullauthor** use the <reg> element, possibly inside a <choice> element, inside either an <author> or <name>

**fullorganization** use the <reg> element, possibly inside a <choice> element, inside a <name type="org">

**howpublished** tag as <note>, possibly using the form <note place="inline">

**institution** used only for issuer of technical reports; tag as <publisher>

**journal** tag as <title level="j"> or <title> within <monogr>

**key** used to specify an alternate sort key for the bibliographic item, for use instead of author's or editor's name; not supported

**meeting** tag as <meeting> or as <note>

**month** use <date>; if the date is not in a trivially parseable form, use the *when* attribute to provide a normalized equivalent in one of the format from XML Schema Part 2: Datatypes Second Edition

**note** tag as <note>

**number** tag as <biblScope unit="issue"> or <biblScope unit="number">; for technical report numbers, use <idno type="docno">

**organization** used only for sponsor of conference; use <name type="org"> within <respStmt> within <meeting> element

**pages** tag as <biblScope unit="pp">

**publisher** tag as <publisher>

**school** used only for institutions at which thesis work is done; tag as <publisher>

**series** tag as <title level="s"> or <title> within <series>

[56] The BibTeX scheme is intentionally compatible with that of Scribe, although it omits some fields used by Scribe. Hence only one list of fields is given here.

**title** tag as <title> in appropriate context or with appropriate *level* value

**volume** tag as <biblScope unit="volume">

**year** tag as <date>; if the date is not in a trivially parseable form, use the *when* attribute to provide an ISO-format equivalent

## 3.13 Passages of Verse or Drama

The following elements are included in the core module for the convenience of those encoding texts which include mixtures of prose, verse and drama.

`<l>` (verse line) contains a single, possibly incomplete, line of verse.

`<lg>` (line group) contains one or more verse lines functioning as a formal unit, e.g. a stanza, refrain, verse paragraph, etc.

`<sp>` (speech) contains an individual speech in a performance text, or a passage presented as such in a prose or verse text.

`<speaker>` contains a specialized form of heading or label, giving the name of one or more speakers in a dramatic text or fragment.

`<stage>` (stage direction) contains any kind of stage direction within a dramatic text or fragment.

Full details of other, more specialized, elements for the encoding of texts which are predominantly verse or drama are described in the appropriate chapter of part three (for verse, see the verse base described in chapter *6. Verse*; for performance texts, see the drama base described in chapter *7. Performance Texts*). In this section, we describe only the elements listed above, all of which can appear in any text, whichever of the three modes prose, verse, or drama may predominate in it.

### 3.13.1 Core Tags for Verse

Like other written texts, verse texts or poems may be hierarchically subdivided, for example into books or cantos. These structural subdivisions should be encoded using the general purpose  or <div1> (etc.) elements described below in chapters *4. Default Text Structure* and *6. Verse.* The fundamental unit of a verse text is the verse line rather than the paragraph, however.

The <l> element is used to mark up verse lines, that is metrical rather than typographic lines. In some modern or free verse, it may be hard to decide whether the typographic line is to be regarded as a verse line or not, but the distinction is quite clear for verse following regular metrical patterns. Where a metrical line is interrupted by the start of a new typographic line, the encoder may choose to ignore the fact entirely or to use the empty <lb> (line beginning) element discussed in *3.11. Reference Systems.* By convention, the start of a metrical line implies the start of a typographic line; hence there is no need to introduce an <lb> tag at the start of every <l> element, but only at places where a new typographic line starts within a metrical line, as in the following example:

```
<l>Of Mans First Disobedience, and<lb/> the Fruit</l>
<l>Of that Forbidden Tree, whose<lb/> mortal tast</l>
<l>Brought Death into the World,<lb/> and all our woe,</l>
<l>With loss of Eden, till one greater Man</l>
<l>Restore us, and regain the blissful Seat...</l>
```

In the original copy text, the presence of an ornamental capital at the start of the poem means that the measure is not wide enough to print the first four lines on four lines; instead each metrical line occupies two typographic lines, with the second beginning at the point indicated. Note that this encoding makes no attempt to preserve information about the whitespace or indentation associated with either kind of line; if regarded as essential, this information would be recorded using the *rend* or *rendition* attributes discussed in *1.3.1.1. Global Attributes.*

The <l> element should not be used to represent typographic lines in non-verse materials: if the lineation of a prose text is considered important for analysis, the beginning of each line should be marked with the <lb> element. Alternatively, a neutral segmentation element such as <seg> or <ab> may be used; see further discussion of these elements in chapter *17. Linking, Segmentation, and Alignment.* The <l> element is a member of the `model.lLike` class, which is a subclass of the `model.divPart` class, along with elements from the `model.pLike` (paragraph-like) class.

In some verse forms, regular groupings of lines are regarded as units of some kind, often identified by a regular verse scheme. In stichic verse and couplets, groups of lines analogous to paragraphs are often indicated by indentation. In other verse forms, lines are grouped into irregular sequences indicated simply by whitespace. The <lg> or line group element may be used to mark any such grouping of elements from the `model.lLike` class. As a member of the `att.typed` class, the <lg> element bears the following attributes:

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

- `@type` characterizes the element in some sense, using any convenient classification scheme or typology.
- `@subtype` (subtype) provides a sub-categorization of the element, if needed.

which may be used to further categorize the line group where this is felt desirable, as in the following example. This example also demonstrates the *rend* attribute to indicate whether or not a line is indented.

```
<lg>
  <l>Come fill up the Glass,</l>
  <l  rend="indent">Round, round let it pass,</l>
  <l>'Till our Reason be lost in our Wine:</l>
  <l  rend="indent">Leave Conscience's Rules</l>
  <l  rend="indent">To Women and Fools,</l>
  <l>This only can make us divine.</l>
</lg>
<lg  n="Chorus"  type="refrain">
  <l>Then a Mohock, a Mohock I'll be,</l>
  <l>No Laws shall restrain</l>
  <l>Our Libertine Reign,</l>
  <l>We'll riot, drink on, and be free.</l>
</lg>
```

For some kinds of analysis, it may be useful to identify different kinds of line group within the same piece of verse. Such line groups may self-nest, in much the same way as the un-numbered  element described in chapter *4. Default Text Structure*. For example:

```
<lg  type="sonnet">
  <lg  type="octet">
    <l>Thus speaks the Muse, and bends her brow severe:—</l>
    <l>“Did I, <name>Lætitia</name>, lend my choicest lays,</l>
    <l>And crown thy youthful head with freshest bays,</l>
    <l>That all the' expectance of thy full-grown year</l>
    <l>Should lie inert and fruitless? O revere</l>
    <l>Those sacred gifts whose meed is deathless praise,</l>
    <l>Whose potent charms the' enraptured soul can raise</l>
    <l>Far from the vapours of this earthly sphere!</l>
  </lg>
  <lg  type="sestet">
    <l>Seize, seize the lyre! resume the lofty strain!</l>
    <l>'T is time, 't is time! hark how the nations round</l>
    <l>With jocund notes of liberty resound,—</l>
    <l>And thy own <name>Corsica</name> has burst her chain!</l>
    <l>O let the song to <name>Britain's</name> shores rebound,</l>
    <l  rend="indent(-1)">Where Freedom's once-loved voice is heard,
        alas! in vain.”</l>
  </lg>
</lg>
```

It is often the case that verse line boundaries conflict with the boundaries of other structural elements. In the following example, the single verse line ‘A Workeman in’t… welcome’ is interrupted by a stage direction:

```
<l>Thou fumblest <name>Eros</name>, and my Queenes a Squire</l>
<l>More tight at this, then thou: Dispatch. O Loue,</l>
<l>That thou couldst see my Warres to day, and knew'st</l>
<l>The Royall Occupation, thou should'st see</l>
<l  part="I">A Workeman in't.</l>
<stage>Enter an Armed Soldier.</stage>
<l  part="F">Good morrow to thee, welcome.</l>
```

In this encoding, the *part* attribute is used, as with , to indicate that the last two <l> elements should be regarded as the initial and final parts of a single line, rather than as two lines.

The same technique may be used where verse lines are collected together into units such as verse paragraphs:

```
<lg  n="6"  type="para">
<!-- ... -->
  <l>Unprofitably travelling toward the grave,</l>
  <l>Like a false steward who hath much received</l>
  <l  part="I">And renders nothing back.</l>
</lg>
<lg  type="para"  n="7">
  <l  part="F">Was it for this</l>
  <l>That one, the fairest of all rivers, loved</l>
  <l>To blend his murmurs with my nurse's song,</l>
<!-- ... -->
</lg>
```

The *part* attribute may also be attached to an <lg> element to indicate that it is incomplete, for example because it forms part of a group that is divided between two speakers, as in the following example:

```
<sp>
  <speaker>First Voice</speaker>
  <lg  type="stanza"  part="I">
    <l>But why drives on that ship so fast</l>
    <l>Withouten wave or wind?</l>
  </lg>
</sp>
<sp>
  <speaker>Second Voice</speaker>
  <lg  type="stanza"  part="F">
    <l>The air is cut away before,</l>
    <l>And closes from behind.</l>
  </lg>
</sp>
```

For alternative methods of aligning groups of lines which do not form simple hierarchic groups, or which are discontinuous, see the more detailed discussion in chapter *17. Linking, Segmentation, and Alignment*. For discussion of other elements and attributes specific to the encoding of verse, see chapter *6. Verse*.

### 3.13.2 Core Tags for Drama

Like other written texts, dramatic and other *performance texts* such as cinema or TV scripts are often hierarchically organized, for example into acts and scenes. These structural subdivisions should be encoded using the general purpose  or <div1> (etc.) elements described below in chapters *4. Default Text Structure* and *7. Performance Texts*. Within these divisions, the body of a performance text typically consists of *speeches*, often prefixed by a phrase indicating who is speaking, and occasionally interspersed with stage directions of various kinds.

In the following simple example, each speech consists of a single paragraph:

```
<div2  n="I.2"  type="scene">
  <head>Scene 2.</head>
  <stage  type="setting">Peachum, Filch.</stage>
  <sp>
    <speaker>FILCH.</speaker>
    <p>Sir, Black Moll hath sent word her Trial comes on in
        the Afternoon, and she hopes you will order Matters
        so as to bring her off.</p>
  </sp>
  <sp>
    <speaker>PEACHUM.</speaker>
    <p>Why, she may plead her Belly at worst; to my
        Knowledge she hath taken care of that Security.
        But, as the Wench is very active and industrious,
        you may satisfy her that I'll soften the Evidence.</p>
  </sp>
  <sp>
    <speaker>FILCH.</speaker>
    <p>Tom Gagg, sir, is found guilty.</p>
  </sp>
</div2>
```

In the following example, each speech consists of a sequence of verse lines, some of them being marked as metrically incomplete:

```
<div1  n="I"  type="Act">
  <head>ACT I</head>
  <div2  n="1"  type="Scene">
    <head>SCENE I</head>
    <stage  rend="italic">Enter Barnardo and Francisco,
        two Sentinels, at several doors</stage>
    <sp>
      <speaker>Barn</speaker>
      <l  part="Y">Who's there?</l>
    </sp>
    <sp>
      <speaker>Fran</speaker>
      <l>Nay, answer me. Stand and unfold yourself.</l>
    </sp>
    <sp>
      <speaker>Barn</speaker>
      <l  part="I">Long live the King!</l>
    </sp>
    <sp>
      <speaker>Fran</speaker>
      <l  part="M">Barnardo?</l>
    </sp>
    <sp>
      <speaker>Barn</speaker>
      <l  part="F">He.</l>
    </sp>
    <sp>
      <speaker>Fran</speaker>
      <l>You come most carefully upon your hour.</l>
    </sp>
    <sp>
```

```
    <speaker>Barn</speaker>
    <l>'Tis now struck twelve. Get thee to bed, Francisco.</l>
   </sp>
   <sp>
    <speaker>Fran</speaker>
    <l>For this relief much thanks. 'Tis bitter cold,</l>
    <l  part="I">And I am sick at heart.</l>
   </sp>
  </div2>
</div1>
```

In some cases, as here in the First Quarto of *Hamlet*, the printed speaker attributions need to be supplemented by use of the *who* attribute; again, the lines are marked as complete or incomplete:

```
<stage>Enter two Centinels.
<add  place="margin">Now call'd <name  xml:id="barnardo">Bernardo</name> &
  <name  xml:id="francisco">Francesco</name>.</add>
</stage>
<sp  who="#francisco">
  <speaker>1.</speaker>
  <l  part="Y">Stand: who is that?</l>
</sp>
<sp  who="#barnardo">
  <speaker>2.</speaker>
  <l  part="Y">Tis I.</l>
</sp>
<sp  who="#francisco">
  <speaker>1.</speaker>
  <l>O you come most carefully vpon your watch,</l>
</sp>
<sp  who="#barnardo">
  <speaker>2.</speaker>
  <l>And if you meete Marcellus and Horatio,</l>
  <l>The partners of my watch, bid them make haste.</l>
</sp>
<sp  who="#francisco">
  <speaker>1.</speaker>
  <l  part="Y">I will: See who goes there.</l>
</sp>
<stage>Enter Horatio and Marcellus.</stage>
```

By contrast with the preceding examples, the following encodes an early printed edition without making any assumption about which parts are prose or verse:

```
<div1  n="I"  type="act">
  <div2  n="1"  type="scene">
    <head  rend="italic">Actus primus, Scena prima.</head>
    <stage  rend="italic"  type="setting">A tempestuous
        noise of Thunder and Lightning heard: Enter
        a Ship-master, and a Boteswaine.</stage>
    <sp>
      <speaker>Master.</speaker>
      <ab>Bote-swaine.</ab>
    </sp>
```

```
    <sp>
      <speaker>Botes.</speaker>
      <ab>Heere Master: What cheere?</ab>
    </sp>
    <sp>
      <speaker>Mast.</speaker>
      <ab>Good: Speake to th' Mariners: fall
            too't, yarely, or we run our selues a ground,
            bestirre, bestirre. <stage  type="move">Exit.</stage>
      </ab>
    </sp>
    <stage  type="move">Enter Mariners.</stage>
    <sp>
      <speaker>Botes.</speaker>
      <ab>Heigh my hearts, cheerely, cheerely my harts: yare,
            yare: Take in the toppe-sale: Tend to th' Masters whistle:
            Blow till thou burst thy winde, if roome e-nough.</ab>
    </sp>
  </div2>
</div1>
```

The <sp> and <stage> elements should also be used to mark parts of a text otherwise in prose which are presented as if they were dialogue in a play. The following example is taken from a 19th century novel in which passages of narrative and passages of dialogue are mixed within the same chapter:

```
<sp>
  <speaker>The reverend Doctor Opimian</speaker>
  <p>I do not think I have named a single unpresentable fish.</p>
</sp>
<sp>
  <speaker>Mr Gryll</speaker>
  <p>Bream, Doctor: there is not much to be said for bream.</p>
</sp>
<sp>
  <speaker>The Reverend Doctor Opimian</speaker>
  <p>On the contrary, sir, I think there is much to be said for him.
     In the first place ...</p>
  <p>Fish, Miss Gryll — I could discourse to you on fish by the
     hour: but for the present I will forbear ...</p>
</sp>
```

```
<sp>
  <speaker>Lord Curryfin</speaker>
  <stage>(after a pause).</stage>
  <p>
    <q>Mass</q> as the second grave-digger says
     in <title>Hamlet</title>, <q>I cannot tell.</q>
  </p>
</sp>
<p>A chorus of laughter dissolved the sitting.</p>
```

## 3.14 Overview of the Core Module

All the elements described in this chapter are provided by the `core` module.

**Module core: Elements common to all TEI documents**

- *Elements defined*: abbr add addrLine address analytic author bibl biblScope biblStruct binaryObject cb choice cit citedRange corr date del desc distinct divGen editor ellipsis email emph expan foreign gap gb gloss graphic head headItem headLabel hi imprint index item l label lb lg list listBibl measure measureGrp media meeting mentioned milestone monogr name note noteGrp num orig p pb postBox postCode ptr pubPlace publisher q quote rb ref reg relatedItem resp respStmt rs rt ruby said series sic soCalled sp speaker stage street teiCorpus term textLang time title unclear unit
- *Classes defined*: att.milestoneUnit

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 4 Default Text Structure

This chapter describes the default high-level structure for TEI documents. A full TEI document combines metadata describing it, represented by a <teiHeader> element, with the document itself, represented by one or more <text> elements or other elements taken from the `model.resource` class. That is, the <TEI> element is used to group together metadata about an encoded resource (in <teiHeader>, specified by the `header` module, which is fully described in chapter *2. The TEI Header*) with an encoded resource. Possible encoded resources are

- a logical transcription of a source document in a <text> element; the <text> element is specified along with its high-level constituents in the `textstructure` module and described in the remainder of the current chapter
- a diplomatic transcription of a source document in a <sourceDoc> element, which is specified in the `transcr` module and described in chapter *12. Representation of Primary Sources*
- an encoded representation of a text-bearing object as images in a <facsimile> element, which is also specified in the `transcr` module and described in chapter *12. Representation of Primary Sources*
- a collection of contextual information or annotations that provides more detail about another encoded resource (whether in the same or a different TEI document) in a <standOff> element, which is specified in the `linking` module and described in section *17.10. The standOff Container*
- a feature system declaration which can be used to declare the use of <fs> elements in the rest of the document, which is specified in the `iso-fs` module and described in section *19.11. Feature System Declaration*

In a case in which more than one resource related to the same source document share the same metadata, they may be grouped together in a <TEI> element following a single <teiHeader>.

Because the <TEI> can be a child of itself, a set or collection of documents may be represented by an outermost <TEI> element that contains a <teiHeader> with metadata that is applicable to the entire set or collection of transcriptions, and then a complete <TEI> element for each document in the collection or set; each of these <TEI> elements contains a <teiHeader> with metadata that is applicable to the individual document, and one or more <text> or other elements taken from the `model.resource` class.

A variant on this basic form, the <teiCorpus>, is also defined for the representation of language corpora, or other collections of encoded texts. A <teiCorpus> consists of its own metadata in a <teiHeader>, followed by one or more complete <TEI> elements, each combining a <teiHeader> with one or more elements from the `model.resource` class. This permits the encoder to distinguish metadata applicable to the whole collection of encoded texts, which is represented by the outermost <teiHeader>, from that applicable to each of the individual <TEI> elements within the corpus. Further information about the organization and encoding of language corpora is given in chapter *16. Language Corpora*.

Alternatively, the corpus may be represented with a <TEI> element (perhaps with a *type* of corpus) in the same manner as a <teiCorpus>.

In summary, when the default structure module is included in a schema, the following elements are available for the representation of the outermost structure of a TEI document:

`<TEI>` (TEI document) contains a single TEI-conformant document, combining a single TEI header with one or more members of the `model.resource` class. Multiple <TEI> elements may be combined within a <TEI> (or <teiCorpus>) element.

`@version` specifies the version number of the TEI Guidelines against which this document is valid.

`<teiCorpus>` (TEI corpus) contains the whole of a TEI encoded corpus, comprising a single corpus header and one or more <TEI> elements, each containing a single text header and a text.

`<teiHeader>` (TEI header) supplies descriptive and declarative metadata associated with a digital resource or set of resources.

`<text>` (text) contains a single text of any kind, whether unitary or composite, for example a poem or drama, a collection of essays, a novel, a dictionary, or a corpus sample.

As noted above, the <teiHeader> element is formally declared in the `header` module (see chapter *2. The TEI Header*). A TEI document may also contain elements from the `model.resource` class (such as a collection of facsimile images, or a feature system declaration) if the appropriate module is included in a schema (see further *12.1. Digital Facsimiles* and *19.11. Feature System Declaration* respectively). By default, however, this class is not populated and hence only the elements <TEI>, <text>, and <teiCorpus> are available as major parts of a TEI document. These three elements are provided by the `textstructure` module described by the present chapter.

TEI texts may be regarded either as *unitary*, that is, forming an organic whole, or as *composite,* that is, consisting of several components which are in some important sense independent of each other. The distinction is not always entirely obvious: for example a collection of essays might be regarded as a single item in some circumstances, or as a number of distinct items in others. In such borderline cases, the encoder must choose whether to treat the text as unitary or composite; each may have advantages and disadvantages in a given situation.

Whether unitary or composite, the text is marked with the <text> tag and may contain front matter, a text body, and back matter. In unitary texts, the text body is tagged <body>; in composite texts, where the text body consists of a series of subordinate texts or groups, it is tagged <group>. The overall structure of any text, unitary or composite, is thus defined by the following elements:

`<front>` (front matter) contains any prefatory matter (headers, abstracts, title page, prefaces, dedications, etc.) found at the start of a document, before the main body.

`<body>` (text body) contains the whole body of a single unitary text, excluding any front or back matter.

`<group>` (group) contains the body of a composite text, grouping together a sequence of distinct texts (or groups of such texts) which are regarded as a unit for some purpose, for example the collected works of an author, a sequence of prose essays, etc.

`<back>` (back matter) contains any appendixes, etc. following the main part of a text.

The overall structure of a unitary text is:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!-- ... -->
  </teiHeader>
  <text>
    <front>
<!-- front matter of copy text, if any, goes here -->
    </front>
    <body>
<!-- body of copy text goes here -->
    </body>
    <back>
<!-- back matter of copy text, if any, goes here -->
    </back>
  </text>
</TEI>
```

The overall structure of a composite text made up of two unitary texts is:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!-- ... -->
  </teiHeader>
  <text>
    <front>
<!-- front matter for composite text -->
    </front>
    <group>
      <text>
        <front>
<!-- front matter of first unitary text, if any -->
        </front>
        <body>
<!-- body of first unitary text -->
        </body>
        <back>
```

```
<!-- back matter of first unitary text, if any -->
      </back>
    </text>
    <text>
      <body>
<!-- body of second unitary text -->
      </body>
    </text>
  </group>
  <back>
<!-- back matter for composite text, if any -->
  </back>
 </text>
</TEI>
```

Finally, a <floatingText> element is provided for the case where one text is embedded within another, but does not contribute to its hierarchical organization, for example because it interrupts it, or simply quoted within it. This is useful in such common literary contexts as the 'play within a play' or the narrative interrupted by other (often deeply nested) multiple narratives.

Each of these elements is further described in the remainder of this chapter. Elements <front> and <back> are further discussed in sections *4.5. Front Matter* and *4.7. Back Matter*. The <group> and <floatingText> elements, used for more complex or composite text structures, are further discussed in section *4.3. Grouped and Floating Texts*. Other textual elements, such as paragraphs, lists or phrases, which nest within these major structural elements, are discussed in chapter *3. Elements Available in All TEI Documents*, in the case of elements which can appear in any kind of document, or elsewhere in the case of elements specific to particular kinds of document.

## 4.1 Divisions of the Body

In some texts, the body consists simply of a sequence of low-level structural items, referred to here as *components* or *component-level elements* (see section *1.3. The TEI Class System*). Examples in prose texts include paragraphs or lists; in dramatic texts, speeches and stage directions; in dictionaries, dictionary entries. In other cases sequences of such elements will be grouped together hierarchically into textual divisions and subdivisions, such as chapters or sections. The names used for these structural subdivisions of texts vary with the genre and period of the text, or even at the whim of the author, editor, or publisher. For example, a major subdivision of an epic or of the Bible is generally called a 'book', that of a report is usually called a 'part' or 'section', that of a novel a 'chapter'—unless it is an epistolary novel, in which case it may be called a 'letter'. Even texts which are not organized as linear prose narratives, or not as narratives at all, will frequently be subdivided in a similar way: a drama into 'acts' and 'scenes'; a reference book into 'sections'; a diary or day book into 'entries'; a newspaper into 'issues' and 'sections', and so forth.

Because of this variety, these Guidelines propose that all such textual divisions be regarded as occurrences of the same neutrally named elements, with an attribute *type* used to categorize elements independently of their hierarchic level. Two alternative styles are provided for the marking of these neutral divisions: *numbered* and *un-numbered*. Numbered divisions are named <div1>, <div2>, etc., where the number indicates the depth of this particular division within the hierarchy, the largest such division being 'div1', any subdivision within it being 'div2', any further sub-sub-division being 'div3' and so on. Un-numbered divisions are simply named , and allowed to nest recursively to indicate their hierarchic depth. The two styles must *not* be combined within a single <front>, <body>, or <back> element.

### 4.1.1 Un-numbered Divisions

The following element is used to identify textual subdivisions in the un-numbered style:

`<div>` (text division) contains a subdivision of the front, body, or back of a text.

As a member of the class `att.typed`, this element has the following additional attributes:

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

- `@type` characterizes the element in some sense, using any convenient classification scheme or typology.

@*subtype* (subtype) provides a sub-categorization of the element, if needed.

Using this style, the body of a text containing two parts, each composed of two chapters, might be represented as follows:

```
<body>
  <div type="part" n="1">
    <div type="chapter" n="1">
<!-- text of part 1, chapter 1 -->
    </div>
    <div type="chapter" n="2">
<!-- text of part 1, chapter 2 -->
    </div>
  </div>
  <div type="part" n="2">
    <div n="1" type="chapter">
<!-- text of part 2, chapter 1 -->
    </div>
    <div n="2" type="chapter">
<!-- text of part 2, chapter 2 -->
    </div>
  </div>
</body>
```

### 4.1.2 Numbered Divisions

The following elements are used to identify textual subdivisions in the numbered style:

<div1> (level-1 text division) contains a first-level subdivision of the front, body, or back of a text.

<div2> (level-2 text division) contains a second-level subdivision of the front, body, or back of a text.

<div3> (level-3 text division) contains a third-level subdivision of the front, body, or back of a text.

<div4> (level-4 text division) contains a fourth-level subdivision of the front, body, or back of a text.

<div5> (level-5 text division) contains a fifth-level subdivision of the front, body, or back of a text.

<div6> (level-6 text division) contains a sixth-level subdivision of the front, body, or back of a text.

<div7> (level-7 text division) contains the smallest possible subdivision of the front, body or back of a text, larger than a paragraph.

As members of the class att.typed these elements all bear the following additional attributes:

att.typed provides attributes that can be used to classify or subclassify elements in any way.

@*type* characterizes the element in some sense, using any convenient classification scheme or typology.

@*subtype* (subtype) provides a sub-categorization of the element, if needed.

The largest possible subdivision of the body is <div1> element and the smallest possible <div7>. If numbered divisions are in use, a division at any one level (say, <div3>), may contain only numbered divisions at the next lowest level (in this case, <div4>).

Using this style, the body of a text containing two parts, each composed of two chapters, might be represented as follows:

```
<body>
  <div1 type="part" n="1">
    <div2 type="chapter" n="1">
<!-- text of part 1, chapter 1 -->
    </div2>
    <div2 type="chapter" n="2">
<!-- text of part 1, chapter 2 -->
```

```
    </div2>
  </div1>
  <div1 type="part" n="2">
    <div2 n="1" type="chapter">
<!-- text of part 2, chapter 1 -->
    </div2>
    <div2 n="2" type="chapter">
<!-- text of part 2, chapter 2 -->
    </div2>
  </div1>
</body>
```

### 4.1.3 Numbered or Un-numbered?

Within the same <front>, <body>, or <back> element, all hierarchic subdivisions must be marked using either nested  elements, or <div1>, <div2> etc. elements nested as appropriate; the two styles must *not* be mixed.

The choice between numbered and un-numbered divisions will depend to some extent on the complexity of the material: un-numbered divisions allow for an arbitrary depth of nesting, while numbered divisions limit the depth of the tree which can be constructed. Where divisions at different levels should be processed differently (for example to ensure that chapters, but not sections, begin on a new page), numbered divisions slightly simplify the task of defining the desired processing for each level, though this distinction could also be made by supplying this information on the *type* attribute of an un-numbered . Some software may find numbered divisions easier to process, as there is no need to maintain knowledge of the whole document structure in order to know the level at which a division occurs; such software may, however, find it difficult to cope with some other aspects of the TEI scheme. On the other hand, in a collection of many works it may prove difficult or impossible to ensure that the same numbered division always corresponds with the same type of textual feature: a 'chapter' may be at level 1 in one work and level 3 in another.

Whichever style is used, the global *n* and *xml:id* attributes (section *1.3.1.1. Global Attributes*) may be used to provide reference strings or labels for each division of a text, where appropriate. Such labels should be provided for each section which is regarded as significant for referencing purposes (on reference systems, see further section *3.11. Reference Systems*).

As indicated above, the *type* and *subtype* attributes provided by the `att.typed` class may be used to provide a name or description for the division. Typical values might be 'book', 'chapter', 'section', 'part', or (for verse texts) 'book', 'canto', 'stanza', or (for dramatic texts) 'act', 'scene'. The following extended example uses numbered divisions to indicate the structure of a novel, and illustrates the use of the attributes discussed above. It also uses some elements discussed in section *4.2. Elements Common to All Divisions* and the <p> element discussed in section *3.1. Paragraphs*.

```
<div1 type="book" n="I" xml:id="JA0100">
  <head>Book I.</head>
  <div2 type="chapter" n="1" xml:id="JA0101">
    <head>Of writing lives in general, and particularly of Pamela, with a word
        by the bye of Colley Cibber and others.</head>
    <p>It is a trite but true observation, that examples work more forcibly on
        the mind than precepts: ... </p>
<!-- remainder of chapter 1 here -->
  </div2>
  <div2 type="chapter" n="2" xml:id="JA0102">
    <head>Of Mr. Joseph Andrews, his birth, parentage, education, and great
        endowments; with a word or two concerning ancestors.</head>
    <p>Mr. Joseph Andrews, the hero of our ensuing history, was esteemed to
        be the only son of Gaffar and Gammar Andrews, and brother to the
        illustrious Pamela, whose virtue is at present so famous ... </p>
<!-- remainder of chapter 2 here -->
  </div2>
```

```
<!-- remaining chapters of Book 1 here -->
  <trailer>The end of the first Book</trailer>
</div1>
<div1  type="book"  n="II"  xml:id="JA0200">
  <head>Book II</head>
  <div2  type="chapter"  n="1"  xml:id="JA0201">
    <head>Of divisions in authors</head>
    <p>There are certain mysteries or secrets in all trades, from the highest
        to the lowest, from that of <term>prime-ministering</term>, to this of
    <term>authoring</term>, which are seldom discovered unless to members of
        the same calling ... </p>
    <p>I will dismiss this chapter with the following observation: that it
        becomes an author generally to divide a book, as it does a butcher to
        joint his meat, for such assistance is of great help to both the reader
        and the carver. And now having indulged myself a little I will endeavour
        to indulge the curiosity of my reader, who is no doubt impatient to know
        what he will find in the subsequent chapters of this book.</p>
  </div2>
  <div2  type="chapter"  n="2"  xml:id="JA0202">
    <head>A surprising instance of Mr. Adams's short memory, with the
        unfortunate consequences which it brought on Joseph.
    </head>
    <p>Mr. Adams and Joseph were now ready to depart different ways ... </p>
  </div2>
</div1>
```

As an alternative (or complement) to this use of the *type* attribute to characterize neutrally named division elements, the modification mechanisms discussed in section *24.3. Customization* may be used to define new elements such as `<chapter>`, `<part>`, etc. To make this simpler, a single member model class is defined for each of the neutrally named division elements: `model.divLike` (containing ), `model.div1Like` (containing <div1>), `model.div2Like` (containing <div2>), etc. For example, suppose that the body of a text consists of a series of diary entries, each of which is potentially divided into entries for the morning and the afternoon. This might be represented in any of the following ways. First, using the un-numbered style:

```
<body>
  <div  type="entry"  n="1">
    <div  type="morning"  n="1.1">
      <p>[...]</p>
    </div>
    <div  type="afternoon"  n="1.2">
      <p>[...]</p>
    </div>
  </div>
  <div  type="entry"  n="2">
    <div  type="morning"  n="2.1">
      <p>[...]</p>
    </div>
    <div  type="afternoon"  n="2.2">
      <p>[...]</p>
    </div>
  </div>
<!-- ...-->
</body>
```

Equivalently, using the numbered style:

```
<body>
  <div1 type="entry" n="1">
    <div2 type="morning" n="1.1">
      <p>[...]</p>
    </div2>
    <div2 type="afternoon" n="1.2">
      <p>[...]</p>
    </div2>
  </div1>
  <div1 type="entry" n="2">
    <div2 type="morning" n="2.1">
      <p>[...]</p>
    </div2>
    <div2 type="afternoon" n="2.2">
      <p>[...]</p>
    </div2>
  </div1>
<!-- ...-->
</body>
```

Now, assuming a customization in which a new element <diaryEntry> has been added to the model.divLike class:

```
<body
   xmlns:my="http://www.example.org/ns/nonTEI">
  <my:diaryEntry type="entry" n="1">
    <my:diaryEntry type="morning" n="1.1">
      <p>[...]</p>
    </my:diaryEntry>
    <my:diaryEntry type="afternoon" n="1.2">
      <p>[...]</p>
    </my:diaryEntry>
  </my:diaryEntry>
  <my:diaryEntry type="entry" n="1">
    <my:diaryEntry type="morning" n="1.1">
      <p>[...]</p>
    </my:diaryEntry>
    <my:diaryEntry type="afternoon" n="1.2">
      <p>[...]</p>
    </my:diaryEntry>
  </my:diaryEntry>
<!-- ...-->
</body>
```

And finally, assuming a customization in which three new elements have been added: <diaryEntry> to the model.div1Like class, and <amEntry> and <pmEntry> both to the model.div2Like class:

```
<body
   xmlns:my="http://www.example.org/ns/nonTEI">
  <p>
<!-- ... -->
  </p>
  <my:diaryEntry type="entry" n="1">
    <my:amEntry type="morning" n="1.1">
```

```
      <p>[...]</p>
    </my:amEntry>
    <my:pmEntry type="afternoon" n="1.2">
      <p>[...]</p>
    </my:pmEntry>
  </my:diaryEntry>
  <my:diaryEntry type="entry" n="1">
    <my:amEntry type="morning" n="1.1">
      <p>[...]</p>
    </my:amEntry>
    <my:pmEntry type="afternoon" n="1.1">
      <p>[...]</p>
    </my:pmEntry>
  </my:diaryEntry>
<!-- ... -->
</body>
```

More information about the customization techniques exemplified here is provided in *24.3. Customization.*

### 4.1.4 Partial and Composite Divisions

In most situations, the textual subdivisions marked by  or <div1> (etc.) elements will be both complete and identically organized with reference to the original source. For some purposes however, in particular where dealing with unusually large or unusually small texts, encoders may find it convenient to present as textual divisions sequences of text which are incomplete with reference to the original text, or which are in fact an ad hoc agglomeration of tiny texts. Moreover, in some kinds of texts it is difficult or impossible to determine the order in which individual subdivisions should be combined to form the next higher level of subdivision, as noted below.

To overcome these problems, the following additional attributes are defined for all elements in the `att.divLike` class:

`att.divLike` provides attributes common to all elements which behave in the same way as divisions.

- `@org` (organization) specifies how the content of the division is organized.
- `@sample` indicates whether this division is a sample of the original source and if so, from which part.

`att.fragmentable` provides attributes for representing fragmentation of a structural element, typically as a consequence of some overlapping hierarchy.

- `@part` specifies whether or not its parent element is fragmented in some way, typically by some other overlapping structure: for example a speech which is divided between two or more verse stanzas, a paragraph which is split across a page division, a verse line which is divided between two speakers.

For example, an encoder might choose to transcribe only the first two thousand words of each chapter from a novel. In such a case, each chapter might conveniently be regarded as a partial division, and tagged with a  element in the following form:

```
<div n="xx" sample="initial" part="Y"
  type="chapter">
  <p> ... </p>
</div>
```

where xx represents a number for the chapter, and the *part* attribute takes the value Y to indicate that this division is incomplete in some respect. Other possible values for this attribute indicate whether material has been omitted initially (I), finally (F), or in the middle (M) of the division, while the <gap> element (*3.5.3. Additions, Deletions, and Omissions*) may be used to indicate exactly where material has been omitted:

```
<div  n="xx"  part="M"  type="chapter">
  <p> ... </p>
  <gap  extent="2"  reason="sampling"/>
  <p> ... </p>
</div>
```

The <samplingDecl> element in the TEI header should also be used to record the principles underlying the selection of incomplete samples, as further described in section *2.3.2. The Sampling Declaration*.

The following example demonstrates how a newspaper column composed of very short unrelated snippets may be encoded using these attributes:

```
<div1  type="storylist"  org="composite">
  <head>News in brief</head>
  <div2  type="story">
    <head>Police deny <soCalled>losing</soCalled> bomb</head>
    <p>Scotland Yard yesterday denied claims in the Sunday
        Express that anti-terrorist officers trailing an IRA van
        loaded with explosives in north London had lost track of
        it 10 days ago.</p>
  </div2>
  <div2  type="story">
    <head>Hotel blaze</head>
    <p>Nearly 200 guests were evacuated before dawn
        yesterday after fire broke out at the Scandic
        Crown hotel in the Royal Mile, Edinburgh.</p>
  </div2>
  <div2  type="story">
    <head>Test match split</head>
    <p>Test Match Special next summer will be split
        between Radio 5 and Radio 3, after protests this
        year that it disrupted Radio 3's music schedule.</p>
  </div2>
</div1>
```

The *org* attribute on the <div1> element is used here to indicate that individual stories in this group, marked here as <div2>, are really quite independent of each other, although they are all marked as subdivisions of the whole group. They can be read in any order without affecting the sense of the piece; indeed, in some cases, divisions of this nature are printed in such a way as to make it impossible to determine the order in which they are intended to be read. Individual stories can be added or removed without affecting the existing components.

This method of encoding composite texts as composite divisions has some limitations compared with the more general and powerful mechanisms discussed in section *4.3.1. Grouped Texts*. However, it may be preferable in some circumstances, notably where the individual texts are very small.

## 4.2 Elements Common to All Divisions

The divisions of any kind of text may sometimes begin with a brief heading or descriptive title, with or without a byline, an epigraph or brief quotation, or a salutation such as one finds at the start of a letter. They may also conclude with a brief trailer, byline, postscript, or signature. Many of these (e.g. a byline) may appear either at the start or at the end of a text division proper.

To support this heterogeneity, the TEI architecture defines five classes, all of which are populated by this module:

`model.divTop` groups elements appearing at the beginning of a text division.

| | |
|---|---|
| model.divTopPart | groups elements which can occur only at the beginning of a text division. |
| model.divWrapper | groups elements which can appear at either top or bottom of a textual division. |

`model.divBottom` groups elements appearing at the end of a text division.

| | |
|---|---|
| model.divBottomPart | groups elements which can occur only at the end of a text division. |
| model.divWrapper | groups elements which can appear at either top or bottom of a textual division. |

`model.divTopPart` groups elements which can occur only at the beginning of a text division.

| | |
|---|---|
| model.headLike | groups elements used to provide a title or heading at the start of a text division. |

| | |
|---|---|
| opener | (opener) groups together dateline, byline, salutation, and similar phrases appearing as a preliminary group at the start of a division, especially of a letter. |
| signed | (signature) contains the closing salutation, etc., appended to a foreword, dedicatory epistle, or other division of a text. |

`model.divBottomPart` groups elements which can occur only at the end of a text division.

| | |
|---|---|
| closer | (closer) groups together salutations, datelines, bylines, and similar phrases appearing as a final group at the end of a division, especially of a letter. |
| postscript | contains a postscript, e.g. to a letter. |
| signed | (signature) contains the closing salutation, etc., appended to a foreword, dedicatory epistle, or other division of a text. |
| trailer | contains a closing title or footer appearing at the end of a division of a text. |

`model.divWrapper` groups elements which can appear at either top or bottom of a textual division.

| | |
|---|---|
| argument | (argument) contains a formal list or prose description of the topics addressed by a subdivision of a text. |
| byline | (byline) contains the primary statement of responsibility given for a work on its title page or at the head or end of the work. |
| dateline | (dateline) contains a brief description of the place, date, time, etc. of production of a letter, newspaper story, or other work, prefixed or suffixed to it as a kind of heading or trailer. |
| docAuthor | (document author) contains the name of the author of the document, as given on the title page (often but not always contained in a byline). |
| docDate | (document date) contains the date of a document, as given on a title page or in a dateline. |
| epigraph | (epigraph) contains a quotation, anonymous or attributed, appearing at the start or end of a section or on a title page. |
| meeting | contains the formalized descriptive title for a meeting or conference, for use in a bibliographic description for an item derived from such a meeting, or as a heading or preamble to publications emanating from it. |
| salute | (salutation) contains a salutation or greeting prefixed to a foreword, dedicatory epistle, or other division of a text, or the salutation in the closing of a letter, preface, etc. |

### 4.2.1 Headings and Trailers

The <head> element is used to identify a heading prefixed to the start of any textual division, at any level. A given division may contain more than one such element, as in the following example:

```
<div1 n="Etym">
  <head>Etymology</head>
  <head>(Supplied by a late consumptive usher to a
     grammar school)</head>
  <p>The pale Usher — threadbare in coat, heart,
     body and brain; I see him now. He was ever
     dusting his old lexicons and grammars, ...</p>
</div1>
```

Unlike some other markup schemes, the TEI scheme does *not* require that headings attached to textual subdivisions at different hierarchic levels have different identifiers. All kinds of heading are marked identically using the <head> tag; the type or level of heading intended is implied by the immediate parent of the <head> element, which may for example be a <div1>, <div2>, etc., an un-numbered , or any member of the `model.listLike` class. However, as with  elements, the encoder may choose to extend the `model.headLike` class of which <head> is the sole member to include other such elements if required.

In certain kinds of text (notably newspapers), there may be a need to categorize individual headings within the sequence at the start of a division, for example as 'main' headings, or 'detail' headings: this may readily be done using the *type* or *subtype* attribute. Specific elements are provided for certain kinds of heading-like features, (notably <byline>, <dateline>, and <salute>; see further section *4.2.2. Openers and Closers*), but the *type* or *subtype* attributes must be used to discriminate among other forms of heading. These attributes are provided, as elsewhere, by the `att.typed` attribute class of which the <head> element is a member.

In the following example, taken from a British newspaper, the lead story and its associated headlines have been encoded as a  element, with appropriate `model.divTop` elements attached:

```
<div type="story">
  <head rend="underlined" type="sub">President pledges safeguards for 2,400 British
     troops in Bosnia</head>
  <head rend="scream" type="main">Major agrees to enforced no-fly zone</head>
  <byline>By George Jones, Political Editor, in Washington</byline>
  <p>Greater Western intervention in the conflict in
     former Yugoslavia was pledged by President Bush ...</p>
</div>
```

In older writings, the headings or *incipits* may be longer than in modern works. When heading-like material appears in the middle of a text, the encoder must decide whether or not to treat it as the start of a new division. If the phrase in question appears to be more closely connected with what follows than with what precedes it, then it may be regarded as a heading and tagged as the <head> of a new  element. If it appears to be simply inserted or superimposed—as for example the kind of 'pull quotes' often found in newspapers or magazines, then the <quote>, <q>, or <cit> element may be more appropriate.

The <trailer> element, which can appear at the end of a division only, is used to mark any heading-like feature appearing in this position, as in this example:

```
<div type="book" n="I">
  <head>In the name of Christ here begins the
     first book of the ecclesiastical history of Georgius Florentinus,
     known as Gregory, Bishop of Tours.</head>
  <div>
    <head>Chapter Headings</head>
```

```
    <list>
      <item>
<!-- chapter head -->
      </item>
<!-- further chapter heads omitted -->
    </list>
  </div>
  <div>
    <head>In the name of Christ here begins Book I of the history.</head>
    <p>Proposing as I do ...</p>
    <p>From the Passion of our Lord until the death of Saint Martin four
        hundred and twelve years passed.</p>
    <trailer>Here ends the first Book, which covers five thousand, five
        hundred and ninety-six years from the beginning of the world down
        to the death of Saint Martin.</trailer>
  </div>
</div>
```

### 4.2.2 Openers and Closers

In addition to headings of various kinds, divisions sometimes include more or less formulaic opening or closing passages, typically conveying such information as the name and address of the person to whom the division is addressed, the place or time of its production, a salutation or exhortation to the reader, and so on. Divisions in epistolary form are particularly liable to include such features. Additional elements for the detailed encoding of personal names, dates, and places are provided in chapter *14. Names, Dates, People, and Places*. For simple cases, the following elements should be adequate:

`<byline>` (byline) contains the primary statement of responsibility given for a work on its title page or at the head or end of the work.

`<dateline>` (dateline) contains a brief description of the place, date, time, etc. of production of a letter, newspaper story, or other work, prefixed or suffixed to it as a kind of heading or trailer.

`<salute>` (salutation) contains a salutation or greeting prefixed to a foreword, dedicatory epistle, or other division of a text, or the salutation in the closing of a letter, preface, etc.

`<signed>` (signature) contains the closing salutation, etc., appended to a foreword, dedicatory epistle, or other division of a text.

The <byline> and <dateline> elements are used to encode headings which identify the authorship and provenance of a division. Although the terminology derives from newspaper usage, there is no implication that <dateline> or <byline> elements apply only to newspaper texts. The following example illustrates use of the <dateline> and <signed> elements at the end of the preface to a novel:

```
<div  type="preface">
  <head>To Henry Hope.</head>
  <p>It is not because this volume was conceived and partly
     executed amid the glades and galleries of the Deepdene,
     that I have inscribed it with your name. ... I shall find a
     reflex to their efforts in your own generous spirit and
     enlightened mind.
  </p>
  <closer>
    <signed  xml:lang="el">D.</signed>
    <dateline>Grosvenor Gate, May-Day, 1844</dateline>
  </closer>
</div>
```

Where a sequence of such elements appear together, either at the beginning or end of an element, it may be convenient to group them together using one of the following elements:

`<opener>` (opener) groups together dateline, byline, salutation, and similar phrases appearing as a preliminary group at the start of a division, especially of a letter.

`<closer>` (closer) groups together salutations, datelines, bylines, and similar phrases appearing as a final group at the end of a division, especially of a letter.

The following examples demonstrate the use of the <opener> and <closer> grouping elements:

```
<div type="narrative" n="6">
  <head>Sixth Narrative</head>
  <head>contributed by Sergeant Cuff</head>
  <div type="fragment" n="6.1">
    <opener>
      <dateline>
        <name type="place">Dorking, Surrey,</name>
        <date>July 30th, 1849</date>
      </dateline>
      <salute>To <name>Franklin Blake, Esq.</name> Sir, —</salute>
    </opener>
    <p>I beg to apologize for the delay that has occurred in the
        production of the Report, with which I engaged to furnish you.
        I have waited to make it a complete Report ...</p>
    <closer>
      <salute>I have the honour to remain, dear sir, your
           obedient servant </salute>
      <signed>
        <name>RICHARD CUFF</name> (late sergeant in the
           Detective Force, Scotland Yard, London). </signed>
    </closer>
  </div>
</div>
```

```
<div type="letter" n="14">
  <head>Letter XIV: Miss Clarissa Harlowe to Miss Howe</head>
  <opener>
    <dateline>Thursday evening, March 2.</dateline>
  </opener>
  <p>On Hannah's depositing my long letter ...</p>
  <p>An interruption obliges me to conclude myself
     in some hurry, as well as fright, what I must ever be,</p>
  <closer>
    <salute>Yours more than my own,</salute>
    <signed>Clarissa Harlowe</signed>
  </closer>
</div>
```

For further discussion of the encoding of dates and of names of persons and places, see section *3.6.4. Dates and Times* and chapter *14. Names, Dates, People, and Places*.

### 4.2.3 Arguments, Epigraphs, and Postscripts

The <argument> element may be used to encode the prefatory list of topics sometimes found at the start of a chapter or other division. It is most conveniently encoded as a list, since this allows each item to be distinguished, but may also simply be presented as a paragraph. The following are thus both equally valid ways of encoding the same argument:

```
<div  type="chap"  n="6">
  <argument>
    <p>Kingston — Instructive remarks on early English history
        — Instructive observations on carved oak and life in general
        — Sad case of Stivvings, junior — Musings on antiquity
        — I forget that I am steering — Interesting result
        — Hampton Court Maze — Harris as a guide.</p>
  </argument>
  <p>It was a glorious morning, late spring or early summer, as you
     care to take it ...</p>
</div>
```

```
<div  type="chap"  n="6">
  <argument>
    <list  type="inline">
      <item>Kingston</item>
      <item>Instructive remarks on early English history</item>
      <item>Instructive observations on carved oak and life in
            general</item>
      <item>Sad case of Stivvings, junior</item>
      <item>Musings on antiquity</item>
      <item>I forget that I am steering</item>
      <item>Interesting result</item>
      <item>Hampton Court Maze</item>
      <item>Harris as a guide.</item>
    </list>
  </argument>
  <p>It was a glorious morning, late spring or early summer, as you
     care to take it ...</p>
</div>
```

An *epigraph* is a quotation from some other work, a saying, or a motto, appearing on a title page, or at the start of a division. It may be encoded using the special-purpose <epigraph> element, as in the following example:

```
<titlePage>
  <docAuthor>E. M. Forster</docAuthor>
  <docTitle>
    <titlePart>Howards End</titlePart>
  </docTitle>
  <epigraph>
    <q>Only connect...</q>
  </epigraph>
</titlePage>
```

When an epigraph contains a quotation, this may often be associated with a bibliographic reference. In such cases, it is recommended additionally to group the quotation and its source together using the <cit> element, as in the following example:

```
<div  n="19"  type="chap">
  <head>Chapter 19</head>
  <epigraph>
    <cit>
```

```
     <quote>I pity the man who can travel
           from Dan to Beersheba, and say <q>'Tis all
              barren;</q> and so is all the world to him
           who will not cultivate the fruits it offers.
     </quote>
     <bibl>Sterne: Sentimental Journey.</bibl>
   </cit>
  </epigraph>
  <p>To say that Deronda was romantic would be to
     misrepresent him: but under his calm and somewhat
     self-repressed exterior ...</p>
</div>
```

For discussion of quotations appearing other than as epigraphs refer to section *3.3.3. Quotation*.

A *postscript* is a passage added after the signature of a letter or, less frequently, the main portion of the body of a book, article, or essay. In English a postscript is often abbreviated as *P.S.* or *PS*, and postscripts are often introduced by labels with one of these abbreviations, as in the following example.

```
<div  type="letter">
  <opener>
    <dateline>
      <placeName>Newport</placeName>
      <date  when="1761-05-27">May ye 27th 1761</date>
    </dateline>
    <salute>Gentlemen</salute>
  </opener>
  <p>Capt Stoddard's Business
  <lb/>calling him to Providence, have
  <lb/>got him to look at Hopkins brigantine
  <lb/>& if can agree to Purchase her, shall
  <lb/>be much oblig'd for your further
  <lb/>assistance herein, & will acquiesce with
  <lb/>whatever you & he shall Contract
  <lb/>for — I Thank you for your
  <lb/>
    <unclear>Line</unclear> respecting the brigantine & Beg
  <lb/>leave to Recommend the Bearer
  <lb/>to you for your advice & Friendship
  <lb/>in this matter</p>
  <closer>
    <salute>I am your most humble servant</salute>
    <signed>Joseph Wanton Jr</signed>
  </closer>
  <postscript>
    <label>P.S.</label>
    <p>I have Mollases, Sugar,
    <lb/>Coffee & Rum, which
    <lb/>will Exchange with you
    <lb/>for Candles or Oyl</p>
  </postscript>
</div>
```

### 4.2.4 Content of Textual Divisions

Other than elements from the `model.divWrapper`, `model.divTop`, or `model.divBottom` classes, every textual division (numbered or un-numbered) consists of a sequence of ungrouped `model.common` elements (see

*1.3. The TEI Class System*). The actual elements available will depend on the modules in use; in all cases, at least the component-level structural elements defined in the core will be available (paragraphs, lists, dramatic speeches, verse lines and line groups etc.). If the drama module has been selected, then other component- or phrase- level items specialized for performance texts (for example, cast lists or camera angles) will be available, as defined in chapter *7. Performance Texts*) will be available. If the dictionary module is in use, then dictionary entries, related entries, etc. (as defined in chapter *10. Dictionaries*) will also be available; if the module for transcribed speech is in use, then utterances, pauses, vocals, kinesics, etc., as defined in chapter *8.3. Elements Unique to Spoken Texts* will be available; and so on.

Where a text contains low-level elements from more than one module these may appear at any point; there is no requirement that elements from the same module be kept together.

## 4.3 Grouped and Floating Texts

The <group> element discussed in *4.3.1. Grouped Texts* should be used to represent a collection of independent texts which is to be regarded as a single unit for processing or other purposes. The <floatingText> element discussed in *4.3.2. Floating Texts* should be used to represent an independent text which interrupts the text containing it at any point but after which the surrounding text resumes.

`<group>` (group) contains the body of a composite text, grouping together a sequence of distinct texts (or groups of such texts) which are regarded as a unit for some purpose, for example the collected works of an author, a sequence of prose essays, etc.

`<floatingText>` (floating text) contains a single text of any kind, whether unitary or composite, which interrupts the text containing it at any point and after which the surrounding text resumes.

### 4.3.1 Grouped Texts

Examples of composite texts which should be represented using the <group> element include anthologies and other collections. The presence of common front matter referring to the whole collection, possibly in addition to front matter relating to each individual text, is a good indication that a given text might usefully be encoded in this way; this structure may be found useful in other circumstances too.

For example, the overall structure of a collection of short stories might be encoded as follows:

```
<text>
  <front>
    <docTitle>
      <titlePart> The Adventures of Sherlock Holmes
      </titlePart>
    </docTitle>
    <docImprint>First published in <title>The Strand</title>
        between July 1891 and December 1892</docImprint>
<!-- any other front matter specific to this collection -->
  </front>
  <group>
    <text>
      <front>
        <head  rend="italic">Adventures of Sherlock
               Holmes</head>
        <docTitle>
          <titlePart>Adventure I. —</titlePart>
          <titlePart>A Scandal in Bohemia</titlePart>
        </docTitle>
        <byline>By A. Conan Doyle.</byline>
      </front>
      <body>
        <p>To Sherlock Holmes she is always
        <emph>the</emph> woman. ... </p>
<!-- remainder of A Scandal in Bohemia here -->
      </body>
```

```
    </text>
    <text>
      <front>
        <head  rend="italic">Adventures of Sherlock Holmes</head>
        <docTitle>
          <titlePart>Adventure II. —</titlePart>
          <titlePart>The Red-Headed League</titlePart>
        </docTitle>
        <byline>By A. Conan Doyle.</byline>
      </front>
      <body>
        <p>I had called upon my friend, Mr. Sherlock Holmes, one day
               in the autumn of last year and found him in deep conversation
               with a very stout, florid-faced, elderly gentleman with fiery red hair …
        </p>
<!-- remainder of The Red Headed League here -->
      </body>
    </text>
    <text>
      <front>
        <head  rend="italic">Adventures of Sherlock Holmes</head>
        <docTitle>
          <titlePart>Adventure XII. —</titlePart>
          <titlePart>The Adventure of the Copper Beeches</titlePart>
        </docTitle>
        <byline>By A. Conan Doyle.</byline>
      </front>
      <body>
        <p>
          <q>To the man who loves art for its
                   own sake,</q> remarked Sherlock Holmes ...


<!-- remainder of The Copper Beeches here -->

               ... she is now the head of a private school
               at Walsall, where I believe that she has
               met with considerable success.</p>
      </body>
    </text>
<!-- end of The Copper Beeches -->
  </group>
</text>
<!-- end of the Adventures of Sherlock Holmes -->
```

A text which is a member of a group may itself contain groups. This is quite common in collections of verse, but may happen in any kind of text. As an example, consider the overall structure of a typical collection, such as the *Muses Library* edition of Crashaw's poetry. Following a critical introduction and table of contents, this work contains the following major sections:

- *Steps to the Temple* (a collection of verse as published in 1648)
- *Carmen deo Nostro* (a second collection, published in 1652)
- *The Delights of the Muses* (a third collection, published in 1648)
- *Posthumous Poems,* I (a collection of fragments all taken from a single manuscript)
- *Posthumous Poems,* II (a further collection of fragments, taken from a different manuscript)

The first three of these collections each has a reasonable claim to be considered as a text in its own right, and may therefore be encoded as such. It is rather more arbitrary as to whether the two posthumous collections should be treated as two groups, following the practice of the *Muses Library* edition. An encoder might elect to combine the two into a single group or simply to treat each fragment as an ungrouped unitary text.

The *Muses Library* edition reprints the whole of each of the three original collections, including their original front matter (title pages, dedications etc.). These should be encoded using the <front> element and its constituents (on which see further section *4.5. Front Matter*), while the body of each collection should be encoded as a single <group> element. Each individual poem within the collections should be encoded as a distinct <text> element. The beginning of the whole collection would thus appear as follows (for further discussion of the use of the elements  and <lg> for textual subdivision of verse, see section *3.13.1. Core Tags for Verse* and chapter *6. Verse*):

```
<text>
  <front>
    <titlePage>
      <docTitle>
        <titlePart>The poems of Richard Crashaw</titlePart>
      </docTitle>
      <byline>Edited by J.R. Tutin</byline>
    </titlePage>
    <div  type="preface">
      <head>Editor's Note</head>
      <p>A few words are necessary ... </p>
    </div>
  </front>
  <group>
    <text>
      <front>
        <titlePage>
          <docTitle>
            <titlePart>Steps to the Temple, Sacred Poems</titlePart>
          </docTitle>
        </titlePage>
        <div  type="address">
          <head>The Preface to the Reader</head>
          <p>Learned Reader, The Author's friend will not usurp much
                    upon thy eye ... </p>
        </div>
      </front>
      <group>
        <text>
          <front>
            <docTitle>
              <titlePart>Sospetto D'Herode</titlePart>
            </docTitle>
          </front>
          <body>
            <div1  type="book"  n="Herod I">
              <head>Libro Primo</head>
              <epigraph>
                <l>Casting the times with their strong signs</l>
              </epigraph>
              <lg  n="I.1"  type="stanza">
                <l>Muse! now the servant of soft loves no more</l>
                <l>Hate is thy theme and Herod whose unblest</l>
                <l>Hand (O, what dares not jealous greatness?) tore</l>
                <l>A thousand sweet babes from their mothers' breast,</l>
                <l>The blooms of martyrdom ...</l>
```

```
            </lg>
          </div1>
        </body>
      </text>
      <text>
        <front>
          <docTitle>
            <titlePart>The Tear</titlePart>
          </docTitle>
        </front>
        <body>
          <lg  n="I">
            <l>What bright soft thing is this</l>
            <l>Sweet Mary, thy fair eyes' expense?</l>
          </lg>
        </body>
      </text>
<!-- remaining poems of the Steps to the Temple appear
here, each tagged as a distinct text element -->
     </group>
     <back>
<!-- back matter for the Steps to the Temple -->
     </back>
   </text>
   <text>
<!-- start of Carmen deo Nostro -->
     <front/>
     <group>
       <text/>
       <text/>
<!-- more texts here -->
     </group>
   </text>
   <text>
<!-- start of The Delights of the Muses -->
     <group>
       <text/>
       <text/>
<!-- more texts here -->
     </group>
   </text>
  </group>
  <back>
<!-- back matter for the whole collection -->
  </back>
</text>
```

The <group> element may be used in this way to encode any kind of collection of which the constituents are regarded by the encoder as texts in their own right. Examples include anthologies or collections of verse or prose by multiple authors, florilegia, or commonplace books, journals, day books, etc. As a fairly typical example, we consider *The Norton Book of Travel*, an anthology edited by Paul Fussell and published in 1987 by W. W. Norton. This work comprises the following major sections:

1. Front matter (title page, acknowledgments, introductory essay)
2. The Beginnings
3. The Eighteenth Century and the Grand Tour

4. The Heyday
5. Touristic Tendencies
6. Post Tourism
7. Back matter (permissions list, index)

Each titled section listed above comprises a group of extracts or complete texts from writers of a given historical period, preceded by an introductory essay. For example, the second group listed above contains, inter alia, the following:

1. Prefatory essay
2. Five letters by Lady Mary Wortley Montagu
3. An extract from Swift's *Gulliver's Travels*
4. Two poems by Alexander Pope
5. Two extracts from Boswell's Journal
6. A poem by William Blake

Each group of writings by a single author is preceded by a brief biographical notice. Some of the extracts are quite lengthy, containing several chapters or other divisions; others are quite short. As the above list indicates, the texts included range across all kinds of material: verse, prose, journals and letters.

The easiest way of encoding such an anthology is to treat each individual extract as a text in its own right. A sequence of texts by a single author, together with the biographical note preceding it, can then be treated as a single <group> element within the larger <group> formed by the section. The sequence of single or composite texts making up a single section of the work is likewise treated, together with its prefatory essay, as a single <group> within the work. Schematically:

```
<text>
<!-- the whole anthology -->
  <front>
<!-- title page, acknowledgments, introductory essay -->
  </front>
  <group>
<!-- body of anthology starts here -->
    <group>
      <head>The Beginnings</head>
<!-- sequence of texts or groups -->
    </group>
    <group>
<!-- The Eighteenth Century and the Grand Tour -->
      <text>
<!-- prefatory essay by editor -->
      </text>
      <group>
<!-- Section on Lady Mary Wortley Montagu starts -->
        <text>
<!-- biographical notice by editor -->
        </text>
        <text>
<!-- first letter -->
        </text>
        <text>
<!-- second letter -->
        </text>
<!-- ... -->
      </group>
```

```
<!-- end of Montagu section -->
     <text>
<!-- single text by Jonathan Swift starts -->
       <front>
<!-- biographical notice by editor -->
       </front>
       <body/>
     </text>
<!-- end of Swift section -->
     <group>
<!-- Section on Alexander Pope starts -->
       <text>
<!-- biographical notice by editor -->
       </text>
       <text>
<!-- first poem -->
       </text>
       <text>
<!-- second poem -->
       </text>
     </group>
<!-- end of Pope section -->
<!-- ... -->
   </group>
<!-- end of 18th century section -->
   <group>
     <head>The Heyday</head>
<!-- texts and subgroups -->
   </group>
<!-- ... -->
  </group>
<!-- end of the anthology proper -->
  <back>
<!-- back matter for anthology -->
  </back>
</text>
```

Note that the editor's introductory essays on each author may be treated as texts in their own right (as the essays on Lady Mary Wortley Montagu and Alexander Pope have been treated above), or as front matter to the embedded text, as the essay on Swift has been. The treatment in the example is intentionally inconsistent, to allow comparison of the two approaches. Consistency can be imposed either by treating the Swift section as a <group> containing one text by Swift and one by the editor, or by treating the Montagu and Pope sections as <text> elements containing the editor's essays as front matter. Marked in the second way, the Pope section of the book would look like this:

```
<text>
<!-- Section on Alexander Pope starts -->
  <front>
<!-- biographical notice by editor -->
  </front>
  <group>
    <text>
<!-- first poem -->
    </text>
    <text>
<!-- second poem -->
    </text>
```

```
  </group>
</text>
<!-- end of Pope section-->
```

The essays on 'The Eighteenth Century and the Grand Tour' and other larger sections could also be tagged as 'front' matter in the same way, by treating the larger sections as <text> elements rather than <group> elements.

Where, as in this case, an anthology contains different kinds of text (for example, mixtures of prose and drama, or transcribed speech and dictionary entries, or letters and verse), the elements to be encoded will of course be drawn from more than one module. The elements provided by the core module described in chapter *3. Elements Available in All TEI Documents* should however prove adequate for most simple purposes, where prose, drama, and verse are combined in a single collection.

For anthologies of short extracts such as commonplace books, it may often be preferable to regard each extract not as a text in its own right but simply as a quotation or <cit> element. The following component-level elements may be used to encode quotations of this kind:

`<cit>` (cited quotation) contains a quotation from some other document, together with a bibliographic reference to its source. In a dictionary it may contain an example text with at least one occurrence of the word form, used in the sense being described, or a translation of the headword, or an example.

`<quote>` (quotation) contains a phrase or passage attributed by the narrator or author to some agency external to the text.

For example, the chapter of 'extracts' which appears in the front matter of Melville's *Moby Dick* might be encoded as follows:

```
<div  n="2"  type="chap">
  <head>Extracts</head>
  <head>(Supplied by a sub-sub-Librarian)</head>
  <p>It will be seen that this mere painstaking burrower and
     grubworm of a poor devil of a Sub-Sub appears to have gone
     through the long Vaticans and street-stalls of the earth,
     picking up whatever random allusions to whales he could
     anyways find ...
     Here ye strike but splintered hearts together — there,
     ye shall strike unsplinterable glasses!</p>
  <p>
    <cit>
      <quote>And God created great whales.</quote>
      <bibl>Genesis</bibl>
    </cit>
    <cit>
      <quote>
        <l>Leviathan maketh a path to shine after him;</l>
        <l>One would think the deep to be hoary.</l>
      </quote>
      <bibl>Job</bibl>
    </cit>
    <cit>
      <quote>By art is created that great Leviathan,
            called a Commonwealth or State — (in Latin,
      <mentioned  xml:lang="la">civitas</mentioned>), which
            is but an artificial man.</quote>
      <bibl>Opening sentence of Hobbes's Leviathan</bibl>
    </cit>
  </p>
</div>
```

For more information on the use of the <quote> and <bibl> elements, see sections *3.3.3. Quotation* and *3.12. Bibliographic Citations and References* respectively.

### 4.3.2 Floating Texts

An important characteristic of the unitary or composite text structures discussed so far is that they can be regarded as forming what is mathematically known as a *tesselation* covering the whole of the available text (or text division) at each hierarchic level. Just as an XML document has a single root element containing a single tree, each node of which forms a properly nested sub-tree, so it seems natural to think of the internal structure of a text as decomposable hierarchically into subparts, each of which is a properly nested subtree. While this is undoubtedly true of a large number of documents, it is not true of all. In particular, it is not true of texts which are only partly tesselated at a given level. For example, if a text A is contained by text B in such a way that part of B precedes A and part follows it, we cannot tesselate the whole of B. In such a case, we say that text A is a 'floating' text.

The <floatingText> element is a member of the `model.divPart` class, and can thus appear within any division level element in the same way as a paragraph. For example, texts such as the *Decameron* or the *Arabian Nights* might be regarded as containing many floating texts embedded within another single text, the framing narrative, rather than as groups of discrete texts in which the fragments of framing narrative are regarded as front or back matter.

As an example, we consider an 18th century text *The Lining to the Patch-Work Screen*, by Jane Barker (1726). This lengthy narrative contains nearly a hundred distinct 'tales' embedded (as the title suggests) in a single patchwork. The work begins by introducing the central character, Galecia, but within a few pages launches into a distinct narrative, the story of Captain Manly:

```
<p>Galecia one Evening setting alone in her Chamber by a clear Fire,
and a clean Hearth [...] reflected on the Providence of our
All-wise and Gracious Creator [...] </p>
<p>She was thus ruminating, when a Gentleman enter'd the Room, the
Door being a jar [...] calling for a Candle, she beg'd a thousand
Pardons, engaged him to sit down, and let her know, what had so long
conceal'd him from her Correspondence.
</p>
<pb  n="5"/>
<floatingText>
  <body>
    <head>The Story of <hi>Captain Manly</hi>
    </head>
    <p>Dear Galecia, said he, though you partly know the loose, or rather
        lewd Life that I led in my Youth; yet I can't forbear relating part of
        it to you by way of Abhorrence...

<!-- Captain Manly's story here -->
        I had lost and spent all I had in the World; in which I verified the
        Old Proverb, That a Rolling Stone never gathers Moss,
    </p>
  </body>
</floatingText>
<pb  n="37"/>
```

Following the conclusion of Captain Manly's tale, we are returned to Galecia, and almost immediately after that into two further stories. However, the Galecia narrative returns between each of the texts, which is why we choose to represent them as <floatingText>s:

```
<p>The Gentleman having finish'd his Story, Galecia waited on him to
the Stairs-head; and at her return, casting her Eyes on the Table, she
saw lying there an old dirty rumpled Book, and found in it the
following story: </p>
```

```
<floatingText>
  <body>
    <p> IN the time of the Holy War when
        Christians from all parts went into the Holy Land to oppose the Turks;
        Amongst these there was a certain English Knight...</p>
<!-- rest of story here -->
    <p>The King graciously pardoned the Knight; Richard was kindly receiv'd
        into his Convent, and all things went on in good order: But from hence
        came the Proverb, We must not strike <hi>Robert</hi> for
    <hi>Richard.</hi>
    </p>
  </body>
</floatingText>
<pb  n="43"/>
<p>By this time Galecia's Maid brought up her Supper; after which she
cast her Eyes again on the foresaid little Book, where she found the
following Story, which she read through before she went to bed.
</p>
<floatingText>
  <body>
    <head>The Cause of the Moors Overrunning
    <hi>Spain</hi>
    </head>
    <p>King ——— of Spain at his Death, committed the Government of his
        Kingdom to his Brother Don ——— till his little Son should come of
        Age ...</p>
    <p>Thus the little Story ended, without telling what Misery
        befel the King and Kingdom, by the Moors, who over ran the Country for
        many Years after. To which, we may well apply the Proverb,
    <quote>
       <l>Who drives the Devil's Stages,</l>
       <l>Deserves the Devil's Wages</l>
      </quote>
    </p>
  </body>
</floatingText>
<p>The reading this Trifle of a Story detained Galecia from her Rest
beyond her usual Hour; for she slept so sound the next Morning, that
she did not rise, till a Lady's Footman came to tell her, that his
Lady and another or two were coming to breakfast with her...
</p>
```

In other multi-narrative texts, the individual nested tales may have greater significance than the framing narratives, and it may therefore be preferable to treat the fragments of framing narrative as front or back matter associated with each nested tale. This is commonly done, for example, in texts such as Chaucer's *Canterbury Tales*, where each tale is typically presented with front matter in which the teller of the tale is introduced, and back matter in which the pilgrims comment on it.

It is important to distinguish between the uses of <floatingText> and <quote>. Whereas the semantics of <quote> suggest that its content derives from a source external to the current text, <floatingText> carries no such implication and is simply used whenever the richer content model that it provides is required to support the markup of a part of a text that is presented as a discrete 'inclusion.' In some cases, such inclusions could be considered external (e.g., enclosures, attachments, etc.); often however, as in the examples above, the included text bears no signs of emanating from outside.

<floatingText> and <quote> may be used in combination. For a text with rich internal structure that is quoted at length, <floatingText> might be used within <quote>. Also, like a unitary text, <floatingText> may include one or more quoted sections, each marked with a <quote> element.

## 4.4 Virtual Divisions

Where the whole of a division can be automatically generated, for example because it is derived from another part of this or another document, an encoder may prefer not to represent it explicitly but instead simply mark its location by means of a processing instruction, or by using the special purpose <divGen> element:

**<divGen>** (automatically generated text division) indicates the location at which a textual division generated automatically by a text-processing application is to appear.

This element is made available by the `model.divGenLike` class of which it is the sole element. The <divGen> element is a member of the `att.typed` class, from which it inherits the *type* and *subtype* attributes. It may appear wherever a  or <div1> (<div2>, etc.) element may appear.

For example, if the table of contents (toc) for a given work is simply derived by copying the first <head> element from each  element in a text, it might be more easily encoded as follows:

```
<divGen type="toc"/>
```

Similarly, in a digital edition combining a transcribed version of some text with a translated version of it, it may be desired to represent the transcript, the translation, and an aligned version of the two as three distinct divisions. This could be achieved by an encoding like the following:

```
<div>
<!-- transcript here-->
</div>
<div>
<!-- translation here -->
</div>
<divGen type="alignment"/>
```

The processing to be carried out when a <divGen> element is rendered will be determined by the application program or stylesheet in use: the function of the TEI markup is simply to identify the location at which the virtual division is to be generated, and also to provide some information about the kind of division to be generated. As such it may be regarded as a special kind of processing instruction, and could equally well be represented by one.

## 4.5 Front Matter

By *front matter* we mean distinct sections of a text (usually, but not necessarily, a printed one), prefixed to it by way of introduction or identification as a part of its production. Features such as title pages or prefaces are clear examples; a less definite case might be the prologue attached to a play. The front matter of an encoded text should not be confused with the TEI header described in chapter *2. The TEI Header*, which serves as a kind of front matter for the computer file itself, not the text it encodes.

An encoder may choose simply to ignore the front matter in a text, if the original presentation of the work is of no interest, or for other reasons; alternatively some or all components of the front matter may be thought worth including with the text as components of the <front> element.[57] With the exception of the title page, (on which see section *4.6. Title Pages*), front matter should be encoded using the same elements as the rest of a text. As with the divisions of the text body, no other specific tags are proposed here for the various kinds of subdivision which may appear within front matter: instead either numbered or un-numbered  elements may be used. The following suggested values[58] for the *type* attribute may be used to distinguish various kinds of division characteristic of front matter:

**preface** A foreword or preface addressed to the reader in which the author or publisher explains the content, purpose, or origin of the text.

[57] This decision should be recorded in the <samplingDecl> element of the header.

[58] As with all lists of 'suggested values' for attributes, it is recommended that software written to handle TEI-conformant texts be prepared to recognize and handle these values when they occur, without limiting the user to the values in this list.

**ack** A formal declaration of acknowledgment by the author in which persons and institutions are thanked for their part in the creation of a text.

**dedication** A formal offering or dedication of a text to one or more persons or institutions by the author.

**abstract** A summary of the content of a text as continuous prose.

**contents** A table of contents, specifying the structure of a work and listing its constituents. The <list> element should be used to mark its structure.

**frontispiece** A pictorial frontispiece, possibly including some text.

The following extended example demonstrates how various parts of the front matter of a text may be encoded. The front part begins with a title page, which is presented in section *4.6. Title Pages* below. This is followed by a dedication and a preface, each of which is encoded as a distinct :

```
<div  type="dedication">
  <p>To my parents, Ida and Max Fish</p>
</div>
<div  type="preface">
  <head>Preface</head>
  <p>The answer this book gives to its title question is <q>there is
        and there isn't</q>.</p>
  <p>Chapters 1–12 have been previously published in the
     following journals and collections:
  <list>
     <item>chapters 1 and 3 in <title>New literary History</title>
     </item>
     <item>chapter 10 in <title>Boundary II</title> (1980)</item>
   </list>.
    I am grateful for permission to reprint.</p>
  <signed>S.F.</signed>
</div>
```

The front matter concludes with another  element, shown in the next example, this time containing a table of contents, which contains a <list> element (as described in section *3.8. Lists*). Note the use of the <ptr> element to provide page-references: the implication here is that the target identifiers supplied (fish1, fish2, etc.) will correspond with identifiers used for the  elements containing chapters of the text itself. (For the <ptr> element, see *3.7. Simple Links and Cross-References*.)

```
<div  type="contents">
  <head>Contents</head>
  <list>
    <item>Introduction, or How I stopped Worrying and Learned to Love
        Interpretation <ptr  target="#fish1"/>
    </item>
    <item>
      <list>
        <head>Part One: Literature in the Reader</head>
        <item  n="1">Literature in the Reader: Affective Stylistics
        <ptr  target="#fish2"/>
        </item>
        <item  n="2">What is Stylistics and Why Are They Saying Such
               Terrible Things About It? <ptr  target="#fish3"/>
        </item>
      </list>
    </item>
  </list>
```

```
</div>
<div  xml:id="fish1">
  <head>Introduction</head>
<!-- ... -->
</div>
<div  xml:id="fish2">
  <head>Literature in the Reader</head>
<!-- ... -->
</div>
<div  xml:id="fish3">
  <head>What is stylistics?</head>
<!-- ... -->
</div>
```

Alternatively, the pointers in the index might link to the page beginnings at which a chapter begins, assuming that these have been included in the markup:

```
<!-- ... --><item  n="1">Literature in the Reader: Affective Stylistics
<ref  target="#fish-p24">24</ref>
</item>
<!-- ... -->
<div  type="chapter">
  <head>Literature in the Reader</head>
  <pb  xml:id="fish-p24"/>
<!-- ... -->
</div>
<!-- ... -->
```

The following example uses numbered divisions to mark up the front matter of a medieval text. Note that in this case no title page in the modern sense occurs; the title is simply given as a heading at the start of the front matter. Note also the use of the *type* attribute on the  elements to indicate document elements comparatively unusual in modern books such as the initial prayer:

```
<front>
  <div1  type="incipit">
    <p>Here bygynniþ a book of contemplacyon, þe whiche
        is clepyd <title>þE CLOWDE OF VNKNOWYNG</title>,
        in þe whiche a soule is onyd wiþ GOD.</p>
  </div1>
  <div1  type="prayer">
    <head>Here biginneþ þe preyer on þe prologe.</head>
    <p>God, unto whom alle hertes ben open, & unto whome alle wille
        spekiþ, & unto whom no priue þing is hid: I beseche
        þee so for to clense þe entent of myn hert wiþ þe
        unspekable 3ift of þi grace, þat I may parfiteliche
        loue þee & worþilich preise þee. Amen.</p>
  </div1>
  <div1  type="preface">
    <head>Here biginneþ þe prolog.</head>
    <p>In þe name of þe Fader & of þe Sone &
        of þe Holy Goost.</p>
    <p>I charge þee & I beseeche þee, wiþ as moche
        power & vertewe as þe bonde of charite is sufficient
        to suffre, what-so-euer þou be þat þis book schalt
```

```
      haue in possession ...</p>
  </div1>
  <div1  type="contents">
    <head>Here biginneþ a table of þe chapitres.</head>
    <list>
      <label>þe first chapitre </label>
      <item>Of foure degrees of Cristen mens leuing; & of þe
            cours of his cleping þat þis book was maad vnto.</item>
      <label>þe secound chapitre</label>
      <item>A schort stering to meeknes & to þe werk of þis
            book</item>
      <label>þe fiue and seuenti chapitre</label>
      <item>Of somme certein tokenes bi þe whiche a man may proue
            whеþer he be clepid of God to worche in þis werk.</item>
    </list>
    <trailer>& here eendeþ þe table of þe chapitres.</trailer>
  </div1>
</front>
```

If, however, the table of contents can be automatically generated from the remainder of the text, it may be preferable simply to mark its presence, either by means of an empty <divGen> element or by using an appropriate processing instruction.

## 4.6 Title Pages

Detailed analysis of the title page and other *preliminaries* of older printed books and manuscripts is of major importance in descriptive bibliography and the cataloguing of printed books; such analysis may require a rather more detailed module than that proposed here. The following elements are suggested as a means of encoding the major features of most title pages:

`<titlePage>` (title page) contains the title page of a text, appearing within the front or back matter.

`<docTitle>` (document title) contains the title of a document, including all its constituents, as given on a title page.

`<titlePart>` (title part) contains a subsection or division of the title of a work, as indicated on a title page.
`@type` (type) specifies the role of this subdivision of the title.

`<argument>` (argument) contains a formal list or prose description of the topics addressed by a subdivision of a text.

`<byline>` (byline) contains the primary statement of responsibility given for a work on its title page or at the head or end of the work.

`<docAuthor>` (document author) contains the name of the author of the document, as given on the title page (often but not always contained in a byline).

`<epigraph>` (epigraph) contains a quotation, anonymous or attributed, appearing at the start or end of a section or on a title page.

`<imprimatur>` (imprimatur) contains a formal statement authorizing the publication of a work, sometimes required to appear on a title page or its verso.

`<docEdition>` (document edition) contains an edition statement as presented on a title page of a document.

`<docImprint>` (document imprint) contains the imprint statement (place and date of publication, publisher name), as given (usually) at the foot of a title page.

`<docDate>` (document date) contains the date of a document, as given on a title page or in a dateline.

`<graphic>` (graphic) indicates the location of a graphic or illustration, either forming part of a text, or providing an image of it.

These elements constitute the `model.titlepagePart` class. Together with the <figure> element described in chapter *15. Tables, Formulæ, Graphics, and Notated Music*, any number of elements from this class can appear within a <titlePage> element. The <figure> element is included so as to enable encoders to record the presence of complex

non-textual material on a title page. For simple cases such as printers' ornaments or illustrations the <graphic> element discussed in section *3.10. Graphics and Other Non-textual Components* should be adequate.

The elements listed above, together with the <head> element, also constitute the class `model.pLike.front`. The elements in this class can appear within a minimal <front> element without any need to group them together and encode a complete title page.

Encoders wishing to add new elements to either class may do so using the methods described in section *24.3. Customization.* Two examples of the use of these elements follow. First, the title page of the work discussed earlier in this section:

```
<front>
  <titlePage>
    <docTitle>
      <titlePart  type="main">Is There a Text in This Class?</titlePart>
      <titlePart  type="sub">The Authority of Interpretive Communities</titlePart>
    </docTitle>
    <docAuthor>Stanley Fish</docAuthor>
    <docImprint>
      <publisher>Harvard University Press</publisher>
      <pubPlace>Cambridge, Massachusetts</pubPlace>
      <pubPlace>London, England</pubPlace>
    </docImprint>
  </titlePage>
</front>
```

Second, a characteristically verbose 17th century example. Note the use of the <lb> tag to mark the line beginnings of the original where necessary:

```
<titlePage>
  <docTitle>
    <titlePart  type="main">THE
    <lb/>Pilgrim's Progress
    <lb/>FROM
    <lb/>THIS WORLD,
    <lb/>TO
    <lb/>That which is to come:</titlePart>
    <titlePart  type="sub">Delivered under the Similitude of a
    <lb/>DREAM</titlePart>
    <titlePart  type="desc">Wherein is Discovered,
    <lb/>The manner of his setting out,
    <lb/>His Dangerous Journey; And safe
    <lb/>Arrival at the Desired Countrey.</titlePart>
  </docTitle>
  <epigraph>
    <cit>
      <quote>I have used Similitudes,</quote>
      <bibl>Hos. 12.10</bibl>
    </cit>
  </epigraph>
  <byline>By <docAuthor>John Bunyan</docAuthor>.</byline>
  <imprimatur>Licensed and Entred according to Order.</imprimatur>
  <docImprint>
    <pubPlace>LONDON,</pubPlace>
     Printed for <name>Nath. Ponder</name>
    <lb/>at the <name>Peacock</name> in the <name>Poultrey</name>
    <lb/>near <name>Cornhil</name>, <docDate>1678</docDate>.
```

```
  </docImprint>
</titlePage>
```

Where, as here, it is considered important to encode salient features of the way a title page was originally rendered, the techniques exemplified in *2.3.4. The Tagging Declaration* may also be useful.

Where title pages are encoded, their physical rendition is often of considerable importance. One approach to this requirement would be to use the <seg> tag, described in chapter *17. Linking, Segmentation, and Alignment*, to segment the typographic content of each part of the title page, and then use the global *rend* attribute to specify its rendition. Another would be to use a module specialized for the description of typographic entities such as pages, lines, rules, etc., bearing special-purpose attributes to describe line-height, leading, degree of kerning, font, etc. Further discussion of these problems is provided in chapter *12. Representation of Primary Sources*.

## 4.7 Back Matter

Conventions vary as to which elements are grouped as back matter and which as front. For example, some books place the table of contents at the front, and others at the back. Even title pages may appear at the back of a book as well as at the front. The content model for <back> and <front> elements are therefore identical.

The following suggested values may be used for the *type* attribute on all division elements, in order to distinguish various kinds of division characteristic of back matter:

**appendix** An ancillary self-contained section of a work, often providing additional but in some sense extra-canonical text.

**glossary** A list of terms associated with definition texts ('glosses'): this should be encoded as a <list type="gloss"> (see section *3.8. Lists*).

**notes** A section in which textual or other kinds of notes are gathered together.

**bibliogr** A list of bibliographic citations: this should be encoded as a <listBibl> (see section *3.12. Bibliographic Citations and References*).

**index** Any form of index to the work.

**colophon** A statement appearing at the end of a book describing the conditions of its physical production.

No additional elements are proposed for the encoding of back matter at present. Some characteristic examples follow; first, an index (for the case in which a printed index is of sufficient interest to merit transcription):

```
<back>
  <div  type="index">
    <head>Index</head>
    <list  type="index">
      <item>Actors, public, paid for the contempt attending
            their profession, <ref>263</ref>
      </item>
      <item>Africa, cause assigned for the barbarous state of
            the interior parts of that continent, <ref>125</ref>
      </item>
      <item>Agriculture
      <list  type="indexentry">
         <item>ancient policy of Europe unfavourable to, <ref>371</ref>
         </item>
         <item>artificers necessary to carry it on, <ref>481</ref>
         </item>
         <item>cattle and tillage mutually improve each other, <ref>325</ref>
         </item>
         <item>wealth arising from more solid than that which proceeds
                  from commerce <ref>520</ref>
```

```
        </item>
       </list>
     </item>
     <item>Alehouses, the number of, not the efficient cause of drunkenness,
<ref>461</ref>
     </item>
   </list>
  </div>
</back>
```

Note that if the pagination in the original source have also been explicitly encoded, and given identifiers, the references to them in the above index can more usefully be recorded as links. For example, assuming that the encoding of page 461 of the original source starts like this:

```
<pb  xml:id="P461"/>
```

then the last item above might be encoded more usefully in either of the following forms:

```
<item>Alehouses, the number of, not
the efficient cause of drunkenness, <ref  target="#P461">461</ref>
</item>
<item>Alehouses, the number of, not the efficient cause of drunkenness,
<ptr  target="#P461"/>
</item>
```

Next, a back-matter division in epistolary form:

```
<back>
  <div  type="letter">
    <head>A letter written to his wife, founde with this booke
        after his death.</head>
    <p>The remembrance of the many wrongs offred thee, and thy
        unreproued vertues, adde greater sorrow to my miserable state,
        than I can utter or thou conceiue. ...
        ... yet trust I in the world to come to find mercie, by the
        merites of my Saiuour to whom I commend thee, and commit
        my soule.</p>
    <signed>Thy repentant husband for his disloyaltie,
    <name>Robert Greene.</name>
    </signed>
    <epigraph  xml:lang="la">
      <p>Faelicem fuisse infaustum</p>
    </epigraph>
    <trailer>FINIS</trailer>
  </div>
</back>
```

And finally, a list of corrigenda and addenda with pseudo-epistolary features:

```
<back>
  <div  type="corrigenda">
```

```
   <head>Addenda</head>
   <salute  xml:lang="la">M. Scriblerus Lectori</salute>
   <p>Once more, gentle reader I appeal unto thee, from the shameful
       ignorance of the Editor, by whom Our own Specimen of
   <name>Virgil</name> hath been mangled in such miserable manner, that
       scarce without tears can we behold it. At the very entrance, Instead
       of <q  xml:lang="grc">προλεγομενα</q>, lo!
   <q  xml:lang="grc">προλεγωμενα</q> with an Omega!
       and in the same line <q  xml:lang="la">consulâs</q> with a circumflex!
       In the next page thou findest <q  xml:lang="la">leviter perlabere</q>,
       which his ignorance took to be the infinitive mood of
   <q  xml:lang="la">perlabor</q> but ought to be
   <q  xml:lang="la">perlabi</q> ... Wipe away all these
       monsters, Reader, with thy quill.</p>
  </div>
</back>
```

## 4.8 Module for Default Text Structure

The module described by the present chapter has the following components:

**Module textstructure: Default text structure**

- *Elements defined*: TEI argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema.*

# 5 Characters, Glyphs, and Writing Modes

Chapter *vi Languages and Character Sets* introduced the fundamental notions of language identification and character representation in an encoded TEI document. In this chapter we discuss some additional issues relating to the way that written language is represented in a TEI document. In sections *5.1. Is Your Journey Really Necessary?* and *5.2. Markup Constructs for Representation of Characters and Glyphs* we introduce markup which may be used to represent and document non-standard characters, that is, written symbols for which no codepoint exists in Unicode. The same markup may be used to annotate existing characters according to their visual or other properties, and thus process them as distinct glyphs (see section *5.3. Annotating Characters*), or to define new characters or glyphs (section *5.4. Adding New Characters*). We also provide recommendations concerning the Unicode Private Use Area (*5.5. How to Use Code Points from the Private Use Area.* Finally, in section *5.6. Writing Modes* we discuss ways of documenting the writing mode used in a source text, that is, the directionality of the script, the orientation of individual characters, and related questions.

## 5.1 Is Your Journey Really Necessary?

Despite the availability of Unicode, text encoders still sometimes find that the published repertoire of available characters is inadequate to their needs. This is particularly the case when dealing with ancient languages, for which encoding standards do not yet exist, or where an encoder wishes to represent variant forms of a character or *glyphs*. The module defined by this chapter provides a mechanism to satisfy that need, while retaining compatibility with standards.

When encoders encounter some graphical unit in a document which is to be represented electronically, the first issue to be resolved should be 'Is this really a different character?' To determine whether a particular graphical unit *is* a character or not, see *vi.2.2 Terminology and Key Concepts*.

If the unit is indeed determined to be a character, the next question should be 'Has this character been encoded already?' In order to determine whether a character has been encoded, encoders should follow the following steps:

1. Check the Unicode web site at `https://www.unicode.org`, in particular the page "Where is my Character?", and the associated character code charts. Alternatively, users can check the latest published version of *The Unicode Standard* (Unicode Consortium (2006)), though the web site is often more up to date than the printed version, and should be checked for preference.

   The pictures ('glyphs') in the Unicode code charts are only meant to be representative, not definitive. If a specific form of an already encoded character is required for a project, refer to the guidelines contained below under Annotating Characters. Remember that your encoded document may be rendered on a system which has different fonts from yours: if the specific form of a character is important to you, then you should document it.

2. Check the Proposed New Characters web page (`https://unicode.org/alloc/Pipeline.html`) to see whether the character is in line for approval.

3. Ask on the Unicode email list (`https://www.unicode.org/consortium/distlist.html`) to see whether a proposal is pending, or to determine whether this character is considered eligible for addition to the Unicode Standard.

Since there are now over 130,000 characters in Unicode, chances are good that what you need is already there, but it might not be easy to find, since it might have a different name in Unicode. Editors working with East Asian writing systems should consult the Unihan Database. Look again, this time at other sites, preferably ones which also provide searches based on scripts and languages. For example `https://www.chise.org` (for CJK characters) or `http://www.eki.ee/letter/` (for non-CJK characters) . Take care, however, that all the properties of what seems to be a relevant character are consistent with those of the character you are looking for. For example, if your character is definitely a digit, but the properties of the best match you can find for it say that it is a letter, you may have a character not yet defined in Unicode.

In general, it is advisable to avoid Unicode characters generally described as presentation forms.[59] However, if the character you are looking for is being used in a notation (rather than as part of the orthography of a language) then it is quite acceptable to select characters from the Mathematical Operators block, provided that they have the appropriate properties (i.e. `So`: Symbol, Other; or `Sm`: Symbol, Math).

[59] Specifically, characters in the Unicode blocks Alphabetic Presentation Forms, Arabic Presentation Forms-A, Arabic Presentation Forms-B, Letterlike Symbols,and Number Forms.

An encoded character may be precomposed or it may be formed from base characters and combining diacritical marks. Either will suffice for a character to be ”found” as an encoded character. If there are several possible Unicode characters to choose amongst, it is good practice to consult other colleagues and practitioners to see whether a consensus has emerged in favour of one or other of them.

If, however, no suitable form of your character seems to exist, the next question will be: ‘Does the graphical unit in question represent a variant form of a known character, or does it represent a completely unencoded character?’ If the character is determined to be missing from the Unicode Standard, it would be helpful to submit the new character for inclusion (see `https://unicode.org/pending/proposals.html`). For assistance on writing or submitting a proposal, potential proposers can contact the UC Berkeley Script Encoding Initiative (`http://linguistics.berkeley.edu/sei/`).

These guidelines will help you proceed once you have identified a given graphical unit as either a variant or an unencoded character. Determining this will require knowledge of the contents of the document that you have. The first case will be called *annotation* of a character, while the second case will be called *adding* of a new character. How to handle graphical units that represent variants will be discussed below (*5.3. Annotating Characters*) while the problem of representing new characters will be dealt with in section *5.4. Adding New Characters*.

While there is some overlap between these requirements, distinct specialized markup constructs have been created for each of these cases. These constructs are presented in section *5.2. Markup Constructs for Representation of Characters and Glyphs* below.

## 5.2 Markup Constructs for Representation of Characters and Glyphs

An XML document can, in principle, contain any defined Unicode character. The standard allows these characters to be represented either directly, using an appropriate encoding (UTF-8 by default), or indirectly by means of a *numeric character reference* (NCR), such as `Ä` (A-umlaut). The encoder can also restrict the range of characters which are represented directly in a document (or part of it) by adding a suitable encoding declaration. For example, if a document begins with the declaration `<?xml encoding="iso-8859-1"?>` any Unicode characters which are not in the ISO-8859-1 character set must be represented by NCRs.

The *gaiji* module defined by this chapter adds a further way of representing specific characters and glyphs in a document. (Gaiji is from Japanese 外字, meaning *external characters*.) This allows the encoder to distinguish characters and glyphs which Unicode regards as identical, to add new nonstandard characters or glyphs, and to represent Unicode characters not available in the document encoding by an alternative means.

The mechanism provided here consists functionally of two parts:

1. an element <g>, which serves as a proxy for new characters or glyphs

2. elements <char> and <glyph>, providing information about such characters or glyphs; these elements are stored in the <charDecl> element in the header.

When the gaiji module is included in a schema, the <charDecl> element is added to the `model.encodingDescPart` class, and the <g> element is added to the phrase class. These elements and their components are documented in the rest of this section.

The Unicode standard defines properties for all the characters it defines in the Unicode Character Database, knowledge of which is usually built into text processing systems. If the character represented by the <g> element does not exist in Unicode at all, its properties are not available. If the character represented is an existing Unicode character, but is not available in the document character set recognized by a given text processing system, it may also be convenient to have access to its properties in the same way. The <char> element makes it possible to store properties for use by such applications in a standard way.

The list of attributes (properties) for characters is modelled on those in the Unicode Character Database, which distinguishes *normative* and *informative* character properties. The Unicode Consortium also maintains a separate set of character properties specific to East Asian characters in the Unihan database which TEI fully supports. Lastly, non-Unicode properties may also be supplied. Since the list of properties will vary with different versions of the Unicode Standard, there may not be an exact correspondence between them and the list of properties defined in these Guidelines.

Usage examples for these elements are given below at *5.3. Annotating Characters* and *5.4. Adding New Characters*. The gaiji module itself is formally defined in section *5.10. Formal Definition* below. It declares the following additional elements:

`<charDecl>` (character declarations) provides information about nonstandard characters and glyphs.

`<g>` (character or glyph) represents a glyph, or a non-standard character.
`@ref` points to a description of the character or glyph intended.

The <charDecl> element is a member of the class `model.encodingDescPart`, and thus becomes available within <encodingDesc> when this module is included in a schema. The <g> element is the only member of the class `model.gLike`: this class is referenced as an alternative to plain text in almost every element which contains plain text, thus permitting the <g> element also to appear at such places when this module is included in a schema.

The following elements may appear within a <charDecl> element:

`<desc>` (description) contains a short description of the purpose, function, or use of its parent element, or when the parent is a documentation element, describes or defines the object being documented.

`<char>` (character) provides descriptive information about a character.

`<glyph>` (character glyph) provides descriptive information about a character glyph.

The <char> and <glyph> elements have similar contents and are used in similar ways, but their functions are different. The <char> element is provided to define a character which is not available in the current document character set, for whatever reason, as stated above. The <glyph> element is used to annotate a character that has already been defined somewhere (either in the document character set, or through a <char> element) by providing a specific glyph that shows how a character appeared in the original document. This is necessary since Unicode code points refer not to a single, specific glyph shape of a character, but rather to a set of glyphs, any of which may be used to render the code point in question; in some cases they can differ considerably.

The <glyph> element is provided for cases where the encoder wants to specify a specific glyph (or family of glyphs) out of all possible glyphs. Unfortunately, due to the way Unicode has been defined, there are cases where several glyphs that logically belong together have been given separate code points, especially in the blocks defining East Asian characters. In such cases, <glyph> elements can also be used to express the view that these apparently distinct characters are to be regarded as instances of the same character (see further *5.3. Annotating Characters*).

The Unicode Standard recommends naming conventions which should be followed strictly where the intention is to annotate an existing Unicode character, and which may also be used as a model when creating new names for characters or glyphs[60]:

Within both <char> and <glyph>, the following elements are available:

`<gloss>` (gloss) identifies a phrase or word used to provide a gloss or definition for some other word or phrase.

`<unicodeProp>` (unicode property) provides a Unicode property for a character (or glyph).

`<unihanProp>` (unihan property) holds the name and value of a normative or informative Unihan character (or glyph) property as part of its attributes.

`<localProp>` (locally defined property) provides a locally defined character (or glyph) property.

`<desc>` (description) contains a short description of the purpose, function, or use of its parent element, or when the parent is a documentation element, describes or defines the object being documented.

`<mapping>` (character mapping) contains one or more characters which are related to the parent character or glyph in some respect, as specified by the *type* attribute.

`<figure>` (figure) groups elements representing or containing graphic information such as an illustration, formula, or figure.

`<note>` (note) contains a note or annotation.

Four of these elements (<gloss>, <desc>, <figure>, and <note>) are defined by other TEI modules, and their usage here is no different from their usage elsewhere. The <figure> element, however, is used here only to link to an image of the character or glyph under discussion, or to contain a representation of it in SVG. The <figure> element may

[60] It should be noted, however, that this naming convention cannot meaningfully be applied to East Asian characters; the typical Unicode descriptions for these characters take the form 'CJK Unified Ideograph `U+4E00`', where `U+4E00` is simply the Unicode code point value of the character in question. In cases where no Unicode code point exists, there is little hope of finding a name that helps to identify the character. Names should therefore be constructed in a way meaningful to local practice, for example by using a reference number from a well-known character dictionary or a project-specific serial number.

contain more than one <graphic> element, for example to provide images with different resolution, or in different formats, or may itself be repeated. As elsewhere, the *mimeType* attribute of <graphic> should be used to specify the format of the image.

The <mapping> element is similar to the standard TEI <equiv> element. While the latter is used to express correspondence relationships between TEI concepts or elements and those in other systems or ontologies, the former is used to express any kind of relationship between the character or glyph under discussion and characters or glyphs defined elsewhere. It may contain any Unicode character, or a <g> element linked to some other <char> or <glyph> element, if, for example, the intention is to express an association between two non-standard characters. The type of association is indicated by the *type* attribute, which may take such values as `exact` for exact equivalences, `uppercase` for uppercase equivalences, `lowercase` for lowercase equivalences, `standard` for standardized forms, and `simplified` for simplified characters, etc., as in the following example:

```
<charDecl>
  <char xml:id="aenl">
    <localProp name="name"
      value="LATIN LETTER ENLARGED SMALL A"/>
    <localProp name="entity" value="aenl"/>
    <mapping type="standard">a</mapping>
  </char>
</charDecl>
```

The mapping element may also be used to represent a mapping of the character or (more likely) glyph under discussion onto a character from the private use area as in this example:

```
<charDecl>
  <glyph xml:id="z103">
    <localProp name="name"
      value="LATIN LETTER Z WITH TWO STROKES"/>
    <mapping type="standard">Z</mapping>
    <mapping type="PUA">U+E304</mapping>
  </glyph>
</charDecl>
```

A more precise documentation of the properties of any character or glyph may be supplied using one of the three 'property' elements: <localProp>, <unicodeProp>, or <unihanProp>; these are described in the next section.

### 5.2.1 Character Properties

The Unicode Standard documents 'ideal' characters, defined by reference to a number of *properties* (or attribute-value pairs) which they are said to possess. For example, a lowercase letter is said to have the value `Ll` for the property `General_Category`. The Standard distinguishes between *normative* properties (i.e. properties which form part of the definition of a given character), and *informative* or *additional* properties which are not normative. It also allows for the addition of new properties, and (in some circumstances) alteration of the values currently assigned to certain properties. When making such modifications, great care should be taken not to override standard informative properties for characters which already exist in the Unicode Standard, as documented in Freytag (2006).

The <unicodeProp>, <unihanProp>, and <localProp> elements allow a TEI encoder to record information about a character or glyph:

`<unicodeProp>` (unicode property) provides a Unicode property for a character (or glyph).
- `@name` specifies the normalized name of a Unicode property.
- `@value` specifies the value of a named Unicode property.

`<unihanProp>` (unihan property) holds the name and value of a normative or informative Unihan character (or glyph) property as part of its attributes.
- `@name` specifies the normalized name of a unicode han database (Unihan) property.
- `@value` specifies the value of a named Unihan property

`<localProp>` (locally defined property) provides a locally defined character (or glyph) property.
- *`@name [att.gaijiProp]`* provides the name of the character or glyph property being defined.
- *`@value [att.gaijiProp]`* provides the value of the character or glyph property being defined.

Where the information concerned relates to a property which has already been identified in the Unicode Standard, use of the appropriate Unicode property name with <unicodeProp> is strongly encouraged. The use of available Unihan property names with <unihanProp> is similarly encouraged. Validation rules for property names according to Unicode conventions are incorporated into the TEI schemas. Where neither of these standards suffices use <localProp>.

The three elements for recording Unicode or locally defined properties belong to the `<att.gaijiProp>` class. This class defines two required attributes for record key-value pairs for character properties:

`att.gaijiProp` provides attributes for defining the properties of non-standard characters or glyphs.
- *`@name`* provides the name of the character or glyph property being defined.
- *`@value`* provides the value of the character or glyph property being defined.

For each property, the encoder must supply both a *name* and a *value*. In cases of boolean properties TEI requires an explict true or false *value* attribute:

```
<unicodeProp  name="Ideographic"
  value="false"/>
```

For convenience, we list here some of the normative character properties and their values. For full information, refer to chapter 4 of *The Unicode Standard*, or the online documentation of the Unicode Character Database.

**General_Category** The general category (described in the Unicode Standard chapter 4 section 5) is an assignment to some major classes and subclasses of characters. Suggested values for this property are listed here:

| | |
|---|---|
| `Lu` | Letter, uppercase |
| `Ll` | Letter, lowercase |
| `Lt` | Letter, titlecase |
| `Lm` | Letter, modifier |
| `Lo` | Letter, other |
| `Mn` | Mark, nonspacing |
| `Mc` | Mark, spacing combining |
| `Me` | Mark, enclosing |
| `Nd` | Number, decimal digit |
| `Nl` | Number, letter |
| `No` | Number, other |
| `Pc` | Punctuation, connector |
| `Pd` | Punctuation, dash |
| `Ps` | Punctuation, open |
| `Pe` | Punctuation, close |
| `Pi` | Punctuation, initial quote |
| `Pf` | Punctuation, final quote |
| `Po` | Punctuation, other |
| `Sm` | Symbol, math |
| `Sc` | Symbol, currency |
| `Sk` | Symbol, modifier |
| `So` | Symbol, other |
| `Zs` | Separator, space |
| `Zl` | Separator, line |
| `Zp` | Separator, paragraph |
| `Cc` | Other, control |

| | |
|---|---|
| `Cf` | Other, format |
| `Cs` | Other, surrogate |
| `Co` | Other, private use |
| `Cn` | Other, not assigned |

**Bidi_Class** This property applies to all Unicode characters. It governs the application of the algorithm for bi-directional behaviour, as further specified in Unicode Annex 9, *The Bidirectional Algorithm*. The following 21 different values are currently defined for this property:

| | |
|---|---|
| `L` | Left-to-Right |
| `R` | Right-to-Left |
| `AL` | Right-to-Left Arabic |
| `EN` | European Number |
| `ES` | European Number Separator |
| `ET` | European Number Terminator |
| `AN` | Arabic Number |
| `CS` | Common Number Separator |
| `NSM` | Nonspacing Mark |
| `BN` | Boundary Neutral |
| `B` | Paragraph Separator |
| `S` | Segment Separator |
| `WS` | Whitespace |
| `ON` | Other Neutrals |
| `LRE` | Left-to-Right Embedding |
| `LRO` | Left-to-Right Override |
| `RLE` | Right-to-Left Embedding |
| `RLO` | Right-to-Left Override |
| `PDF` | Pop Directional Format |
| `LRI` | Left-to-Right Isolate |
| `RLI` | Right-to-Left Isolate |
| `FSI` | First Strong Isolate |
| `PDI` | Pop Directional Isolate |

**Canonical_Combining_Class** This property exists for characters that are not used independently, but in combination with other characters, for example the strokes making up CJK (Chinese, Japanese, and Korean) characters. It records a class for these characters, which is used to determine how they interact typographically. The following values are defined in the Unicode Standard: (see Unicode Character Database: Canonical Combining Class Values); these were taken from version 12.1:

| | |
|---|---|
| `0` | Spacing, split, enclosing, reordrant, and Tibetan subjoined |
| `1` | Overlays and interior |
| `7` | Nuktas |
| `8` | Hiragana/Katakana voicing marks |
| `9` | Viramas |
| `10` | Start of fixed position classes |
| `199` | End of fixed position classes |
| `200` | Below left attached |
| `202` | Below attached |
| `204` | Below right attached |
| `208` | Left attached (reordrant around single base character) |
| `210` | Right attached |
| `212` | Above left attached |
| `214` | Above attached |

| | |
|---|---|
| 216 | Above right attached |
| 218 | Below left |
| 220 | Below |
| 222 | Below right |
| 224 | Left (reordrant around single base character) |
| 226 | Right |
| 228 | Above left |
| 230 | Above |
| 232 | Above right |
| 233 | Double below |
| 234 | Double above |
| 240 | Below (iota subscript) |

**Decomposition_Mapping** This property is defined for characters, which may be decomposed, for example to a canonical form plus a typographic variation of some kind. For such characters the Unicode standard specifies both a decomposition type and a decomposition mapping (i.e. another Unicode character to which this one may be mapped in the way specified by the decomposition type). The following types of mapping are defined in the Unicode Standard:

| | |
|---|---|
| font | A font variant (e.g. a blackletter form) |
| noBreak | A no-break version of a space or hyphen |
| initial | An initial presentation form (Arabic) |
| medial | A medial presentation form (Arabic) |
| final | A final presentation form (Arabic) |
| isolated | An isolated presentation form (Arabic) |
| circle | An encircled form |
| super | A superscript form |
| sub | A subscript form |
| vertical | A vertical layout presentation form |
| wide | A wide (or zenkaku) compatibility character |
| narrow | A narrow (or hankaku) compatibility character |
| small | A small variant form (CNS compatibility) |
| square | A CJK squared font variant |
| fraction | A vulgar fraction form |
| compat | Otherwise-unspecified compatibility character |

**Numeric_Value** This property applies for any character which expresses any kind of numeric value. Its value is the intended value in decimal notation.

**mirrored** The mirrored character property is used to properly render characters such as U+0028, OPENING PARENTHESIS independent of the text direction: it has the value Y (character is mirrored) or N (code is not mirrored).

The Unicode Standard also defines a set of informative (but non-normative) properties for Unicode characters. If encoders wish to provide such properties, they should be included using the Unicode name. If a Unicode name exists for a given character this should always be used, however encoders may also supply locally defined names. To tag a Unicode name, use <unicodeProp name="Name"> (or <unihanProp name="Name">). For names specified elsewhere or specified locally use <localProp>.

## 5.3 Annotating Characters

Annotation of a character becomes necessary when it is desired to distinguish it on the basis of certain aspects (typically, its graphical appearance) only. In a manuscript, for example, where distinctly different forms of the letter *r* can be recognized, it might be useful to distinguish them for analytic purposes, quite distinct from the need to provide an accurate representation of the page. A digital facsimile, particularly one linked to a transcribed and

encoded version of the text, will always provide a superior visual representation (for information on how to link a digital facsimile to a transcribed text see *12.1. Digital Facsimiles*), but cannot be used to support arguments based on the distribution of such different forms. Character annotation as described here provides a solution to this problem.[61]

Assuming that we wish to distinguish the variant glyphs from the standard representation for the character concerned, we will need to define distinct <glyph> elements, one for each of the forms of the letter we wish to distinguish:

```
<charDecl>
  <glyph xml:id="r1">
    <localProp name="name"
      value="LATIN SMALL LETTER R WITH ONE FUNNY STROKE"/>
    <localProp name="entity" value="r1"/>
    <figure>
      <graphic url="r1img.png"/>
    </figure>
  </glyph>
  <glyph xml:id="r2">
    <localProp name="name"
      value="LATIN SMALL LETTER R WITH TWO FUNNY STROKES"/>
    <localProp name="entity" value="r2"/>
    <figure>
      <graphic url="r2img.png"/>
    </figure>
  </glyph>
</charDecl>
```

With these definitions in place, occurrences of these two special *r*s in the text can be represented using the element <g>:

```
<p>Wo<g ref="#r1">r</g>ds in this
manusc<g ref="#r2">r</g>ipt are sometimes
written in a funny way.</p>
```

As can be seen in this example, the <glyph> element pointed to from the <g> element will be interpreted as an annotation on the content of the element <g>. This mechanism can be used to represent common manuscript abbreviations or ligatures, as in the following examples:

```
<p> ... <g ref="#Filig">Fi</g>lthy riches...</p>
<!-- in the charDecl -->
<glyph xml:id="Filig">
  <localProp name="Name"
    value="LATIN UPPER F AND LATIN LOWER I LIGATURE"/>
  <figure>
    <graphic url="Filig.png"/>
  </figure>
</glyph>
```

[61] It should be kept in mind that any kind of text encoding is an abstraction and an interpretation of the text at hand, which will not necessarily be useful in reproducing an exact facsimile of the appearance of a manuscript.

```
<p> ... <abbr>
   <g  ref="#per">per</g>
  </abbr> ardua</p>
<!-- in the charDecl -->
<glyph  xml:id="per">
  <localProp  name="Name"
    value="LATIN ABBREVIATION PER"/>
  <figure>
    <graphic  url="per.png"/>
  </figure>
</glyph>
```

(In fact the Unicode Standard does provide a character to represent the Fi ligature; the encoder may however prefer not to use it in order to simplify other text processing operations, such as indexing).

With this markup in place, it will be possible to write programs to analyze the distribution of the different letters *r* as well as produce more 'faithful' renderings of the original. It will also be possible to produce normalized versions by simply ignoring the annotation pointed to by the element <g>.

For brevity of encoding, it may be preferred to predefine internal entities such as the following:

```
  <!ENTITY r1 '<g ref="#r1">r</g>' >
<!ENTITY r2 '<g ref="#r2">r</g>' >
```

which would enable the same material to be encoded as follows:

```
  <p>Wo&r1;ds in this manusc&r2;ipt are
sometimes written in a funny way.</p>
```

The same technique may be used to represent particular abbreviation marks as well as to represent other characters or glyphs. For example, if we believe that the r-with-one-funny-stroke is being used as an abbreviation for receipt, this might be represented as follows:

```
  <abbr>&r1;</abbr>
```

Note however that this technique employs markup objects to provide a link between a character in the document and some annotation on that character. Therefore, it cannot be used in places where such markup constructs are not allowed, notably in attribute values.

Since the need to use these constructs to annotate or define characters occurs frequently in Chinese, Korean, and Japanese documents, here are some issues that are specific to these documents. There are two slightly different versions of the problem. In the first case, due to the way Unicode is defined, there are occasions when more than one glyph is defined for a character. In such an occasion, one might want to retain the character as used, but add information in a way so that a normalizer (for search or indexing operations) could take advantage of this information. To achieve this, we simply define within a <charDecl> element a <glyph> that has two <mapping> elements, as shown here:

```
<charDecl>
  <glyph  xml:id="u8aaa">
```

```
    <mapping  type="Unicode">說</mapping>
    <mapping  type="standard">說</mapping>
  </glyph>
</charDecl>
```

The first of these <mapping>s, of type Unicode, simply maps our glyph to the code point where Unicode defined it. The other one, of type standard, encodes the fact that in our view, this glyph is a variation of the standard character given in the content of the element. We could then use this <glyph> element's unique identifier u8aaa to refer to it from within a text as follows.

```
<g  ref="#u8aaa">說</g>
```

A slightly different, but related problem occurs when we have multiple variants, none of which has been defined in Unicode. In this case, we need to define one as a new character using <char>, and the others as glyphs using <glyph>.

```
<charDecl>
  <char  xml:id="newchar1">
<!-- more properties here -->
  </char>
  <glyph  xml:id="varofnewchar1">
<!-- more properties here -->
    <mapping  type="standard">
      <g  ref="#newchar1"/>
    </mapping>
  </glyph>
</charDecl>
```

The <char> defines a new character, while the <glyph> element then defines a variant glyph of this newly defined character. Additional properties should be specified in order to make these both identifiable.

## 5.4 Adding New Characters

The creation of additional characters for use in text encoding is quite similar to the annotation of existing characters. The same element <g> is used to provide a link from the character instance in the text to a character definition provided within the <charDecl> element. This character definition takes the form of a <char> element. The element <g> itself will usually be empty, but could contain a code point from the Private Use Area (PUA) of the Unicode Standard, which is an area set aside for the very purpose of privately adding new characters to a document. Recommendations on how to use such PUA characters are given in the following section.

In some circumstances, it may be desirable to provide a single precomposed form of a character that is encoded in Unicode only as a sequence of code points. For example, in Medieval Nordic material, a character looking like a lowercase letter Y with a dot and an acute-accent above it may be encountered so frequently that the encoder wishes to treat it as a single precomposed character with one single coded value. In the transcription concerned, the encoder enters this letter as `&ydotacute;`, which when the transcription is processed can then be expanded in one of three ways, depending on the mapping in force. The entity reference might be translated into the sequence of corresponding Unicode code points or into some locally-defined PUA character (say ``) for local processing only. Both these options have disadvantages; the former loses the fact that the sequence of composed characters is regarded as a single object; the second is not reliably portable. Therefore, the recommended representation is to use the <g> element defined by the module defined in this chapter:

```
<g  ref="#ydotacute"/>
```

. This makes it possible for the encoder to provide useful documentation for the particular character or glyph so referenced:

```
<char xml:id="ydotacute">
  <localProp name="Name"
    value="LATIN SMALL LETTER Y WITH DOT ABOVE AND ACUTE"/>
  <localProp name="entity" value="ydotacute"/>
  <mapping type="composed">&#x0079;&#x0307;&#x0301;</mapping>
  <mapping type="PUA">U+E0A4</mapping>
</char>
```

This definition specifies the mapping between this composed character and the individual Unicode-defined code points which make it up. It also supplies a single locally-defined property ('entity') for the character concerned, the purpose of which is to supply a recommended character entity name for the character.

The composition rules for ideographic characters typically require more complex rules than the &ydotacute; above. For these cases Unicode provides dedicated symbols to capture the composition in Ideographic Description Sequences (IDS). Encoders are strongly encouraged to provide IDS for each variant ideograph in the header component of the gaiji module to faciliated greater human and machine readability of rare or unencoded characters, as in the following example:

```
<glyph xml:id="U507D-var">
<!-- more properties here -->
  <mapping type="IDS">□人□</mapping>
  <mapping type="standard">□</mapping>
</glyph>
```

The composition rules and further examples appear in Chapter 18.2: Ideographic Description Characters of the Unicode Standard. Editors should be aware that different sequences can accurately describe the same character. In the example, the character "人" (U+4EBA) could have been substituted with "亻" (U+4EBB). Local preferences about how sequences are constructed should be documented in the <encodingDesc> of the corresponding TEI header (see *2.3. The Encoding Description*). Additionally, a number of online services, such as CHISE, offer querying and retrieving characters via IDS, which facilitates a greater degree of stability across different applications.

Under certain circumstances, Chinese Han characters can be written within a circle. Rather than considering this as simply an aspect of the rendering, an encoder may wish to treat such circled characters as entirely distinct derived characters. For a given character (say that represented by the numeric-character reference 人) the circled variant might conveniently be represented as

```
<g ref="#U4EBA-circled"/>
```

, which references a definition such as the following:

```
<char xml:id="U4EBA-circled">
  <unicodeProp name="Decomposition_Mapping"
    value="cicle"/>
  <localProp name="Name"
    value="CIRCLED IDEOGRAPH 36"/>
  <localProp name="daikanwa" value="36"/>
  <mapping type="standard"> &#x4EBA;
  </mapping>
  <mapping type="PUA"> &#xE000;
  </mapping>
</char>
```

In this example, the 'circled ideograph' character has been defined with two mappings, and with two properties. The two properties are the Unicode-defined character-decomposition which specifies that this is a circled character, using the appropriate terminology (see *5.2.1. Character Properties* above) and a locally defined property known as 'daikanwa'. The two mappings indicate firstly that the standard form of this character is the character 人, and secondly that the character used to represent this character locally is the PUA character . For convenience of local processing this PUA character may in fact appear as content of the <g> element. In general, however, the <g> element will be empty.

## 5.5 How to Use Code Points from the Private Use Area

The developers of the Unicode Standard have set aside an area of the codespace for the private use of software vendors, user groups, or individuals. As of this writing (Unicode 12.1), there are around 137,000 code points available in this area, which should be enough for most needs. No code point assignments will be made to this area by standard bodies and only some very basic default properties have been assigned (which may be overridden where necessary by the mechanism outlined in this chapter). Therefore, unlike all other code points defined by the Unicode Standard, PUA code points should *not* be used directly in documents intended for blind interchange.

In the two previous examples, we mentioned that the variant characters concerned might well be assigned specific code points from the PUA. This might, for example, facilitate the use of a particular font which displays the desired character at this code point in the local processing environment. Since however this assignment would be valid only on the local site, documents containing such code points are unsuitable for blind interchange. During the process of preparing such documents for interchange, any PUA code points should be replaced by an appropriate use of the <g> element, such as <g ref="#xxxx">, thus associating the character required with the documentation of it provided by the referenced <char> element. The PUA character used during the preparation of the document might be recorded in the <char> element, as shown in the example in *5.4. Adding New Characters*, or retained as content of the <g> element. However, since there is no requirement that the same PUA character be used to represent it at the receiving site, and since it may well be the case that this other site has already made an assignment of some other character to the original PUA code point, it is best practice to remove the locally-defined PUA character. It is to be expected that a further translation into the local processing environment at the receiving site will be necessary to handle such characters, during which variant letters can be converted to hitherto unused code points on the basis of the information provided in the <char> element.

This mechanism is rather weak in cases where DOM trees or parsed XML fragments are exchanged, which may increasingly be the case. The best an application can do here is to treat any occurrence of a PUA character only in the context of the local document and use the properties provided through the <char> element as a handle to the character in other contexts.

In the fullness of time, a character may become standardized, and thus assigned a specific code point outside the PUA. Documents which have been encoded using the mechanism must at the least ensure that this changed code point is recorded within the relevant <char> element; it will however normally be simpler to remove the <char> element and replace all occurrences of <g> elements which reference it by occurrences of the newly coded character.

## 5.6 Writing Modes

The scripts used for writing human languages vary not only in the glyphs they use, but also in the way (or ways) that those glyphs are arranged on the writing surface. For the majority of modern languages, writing is arranged as a series of lines which are to be read from top to bottom. Within each line, individual characters are frequently presented from left to right (English, Russian, Greek), but there are also several widely-used scripts which run right-to-left (Arabic, Hebrew). Writing in which the lines of glyphs are presented vertically and read from right to left is also often encountered, notably in East Asian scripts (Sinitic characters, Japanese Kana, Korean Hangul, Vietnamese chữ nôm). In many cases, a language normally uses the same *writing mode* (we use this term to refer to the orientation of individual glyphs within a line and the order in which glyphs and lines should be read), but there are exceptions in which the same language may appear in different modes, for example either vertically or horizontally. Many East Asian scripts were traditionally written from top to bottom within the line, with their lines sequenced from right to left. Although modern Japanese, Chinese, and Korean are often written horizontally, the traditional vertical writing mode is still widely used. There are also comparatively rare cases of ancient scripts written with lines running left to right, each line being read top to bottom (Ancient Uighur, classical Mongolian and Manchu), or scripts such as Ogham where the writing direction may start from the bottom left and run around the edge of an inscribed object.

When different languages are combined, it is possible that different writing modes will be needed: for example, in Hebrew text, running right to left, sequences of Latin digits still run left to right. When different writing modes are available for the same language, it may be that different glyphs will be preferred when the script is used in different modes. For example, when Japanese is written horizontally, the Unicode character U+3001, the 'ideographic comma', is used in preference to Unicode character U+FE11, the vertical mode comma. This ensures that the comma appears in the correct position relative to the surrounding glyphs. Even for scripts which are usually written in exactly the same way, different writing modes may be encountered in particular contexts; for example when a language using Roman script is embedded within vertically-organized Chinese text, it may sometimes be displayed vertically and sometimes horizontally. The writing mode may also vary in response to layout constraints such as those imposed by a complex table, where column or row labels may be written vertically or diagonally to make the most effective use of available space, just as it may vary in response to the size and shape of the carrier in the case of a monumental inscription.

For many, perhaps most, TEI documents there may be no need to encode the writing mode explicitly, even in so-called "mixed mode" texts containing passages written in languages which use different writing modes. Modern printed texts in most European languages, for instance, may be expected to use left-to-right/top-to-bottom directionality; while Arabic or Hebrew texts are expected to run right-to-left/top-to-bottom. In a TEI document, language and script are explicitly stated in the markup using the attribute *xml:lang*; this indication will usually imply a particular default writing mode. Even where this attribute is not used, passages in different scripts will use different Unicode characters, and will thus imply a particular default writing mode.

Consider the case of an English text containing a few Arabic words:

```
The Arabic term مقلم رصاص means "pencil".
```

A correct TEI encoding might read as follows:

```
<s  xml:lang="en">The Arabic term
<term  xml:lang="ar">مقلم رصاص</term> means "pencil".</s>
```

We might assume that it is the presence of the *xml:lang* attribute with value ar that causes processing software to display the Arabic from right to left, but in fact, this is not the case. The order in which the Arabic characters appear when rendered would be the same, even if the markup were not present:

```
<s>The Arabic term مقلم رصاص means "pencil".</s>
```

This is because Arabic glyphs are always displayed right to left, even when they appear within a left-to-right English sentence. Like most other codepoints in the Unicode standard, they have a specific directionality setting which helps any rendering software determine how they should be ordered. The Latin glyph "a" has a strong left-to-right bidirectionality setting, as do the digits 0 to 9; the Hebrew א (alef) is strongly right-to-left. Of course, some glyphs (common punctuation marks such as the period or comma for example) have weak or neutral settings because they may appear in several contexts.

The Unicode Bidirectional Algorithm (Unicode Consortium, 2017) defines a number of rules enabling software to render sequences of characters which have differing directionality properties in a predictable and reliable way, using only those properties. [62]. It should be remembered however that individual sequences of characters are always stored in a file in the order in which they should be read, irrespective of the order in which the characters making up a sequence should be displayed or rendered. For example, in a RTL language such as Hebrew, the first character in a file will be that which is displayed at the rightmost end of the first line of text.

[62] Because this algorithm may not always give the desired result, Unicode also provides a set of "directional formatting characters" (`https://www.unicode.org/reports/tr9/#Directional_Formatting_Characters`). These additional codepoints can be used to signal to rendering software that a specific directionality setting should be turned on or off. However, in the case of documents encoded in XML, there is generally no need to use such characters, and the W3C advises against it unless markup is unavailable. (`https://www.w3.org/International/questions/qa-bidi-controls`)

An encoder wishing to document or to control the order in which sequences of characters in a TEI document are displayed will usually do so by segmenting the text into sequences presented in the desired order and specifying an appropriate language code for each. In situations where this approach may result in ambiguity or lack of precision, or if the encoder wishes to record directional information explicitly in their encoding, we recommend using the global @style attribute to supply detail about the writing mode applicable to the content of any element. The *style* attribute (discussed in *1.3.1.1.3. Rendition Indicators*) permits use of any formatting language; for these purposes however, we recommend use of CSS, which includes a Writing Modes module [63] which permits direct specification of a number of useful properties associated with writing modes, notably `direction` (`ltr` or `rtl`); `writing-mode` (`horizontal-tb`, `vertical-rl`, or `vertical-lr`); and `text-orientation` (`mixed`, `upright`, `sideways` ...) as well as properties affecting the behaviour of the unicode-bidi (bidirectional) algorithm. We discuss and exemplify how these properties may be used below.

The global TEI *style* attribute applies to the element on which it is specified (and in most cases, its descendants). Rather than specify it on every element, it will often be more efficient to express sets of commonly-used styling rules as <rendition> elements in the <teiHeader> and then point to them using the global *rendition* attribute, as further discussed in *2.3.4.1. Rendition*. Although the CSS specifications are mainly used to provide instructions for software when rendering a digital text, they also provide a useful means of describing the visual properties of a pre-existing document in a formal and standardized way.

The next section presents some examples of how CSS can be used to describe a variety of writing modes. A full description of the appearance of a document will probably include many other properties of course.

## 5.7 Examples of Different Writing Modes

The CSS recommendations provides several properties which can be used to encode aspects of the ”writing mode”. The most useful of these is the property ”writing-mode” which may be used to specify a reading-order for both characters within a single line and lines within a single block of text. The property ”text-orientation” may also used to indicate the orientation of individual characters with respect to the line, and the property ”direction” to determine the reading order of characters within a line only. We give some examples of each below.

### 5.7.1 Vertical Writing Modes

The `writing-mode` property is particularly useful for languages which can be written in different writing modes, such as Chinese and Japanese. Its possible values include `horizontal-tb`, `vertical-rl` and `vertical-lr`. Each value has two components: ‘horizontal’ or ‘vertical’ specifies the inline writing direction, while the second component specifies the direction in which lines in a block, and blocks in a sequence are arranged: from top to bottom (as in most European languages, in which lines and paragraphs are arranged from top to bottom on a page), from right to left (as in the case of Japanese written vertically), or left-to-right (as in the case of Mongolian).

The following example shows three versions of the same poem: first in Japanese, written top to bottom; next in *romaji* (Japanese in Latin script); and finally in an English translation.

We might encode this as follows:

```
<div>
  <lg  xml:lang="ja"
    style="writing-mode: vertical-rl">
    <l>古池や</l>
    <l>蛙</l>
    <l>飛び込む</l>
    <l>水の音</l>
  </lg>
  <lg  xml:lang="ja-Latn"
    style="writing-mode: horizontal-tb">
    <l>furu ike ya</l>
    <l>kawazu tobikomu</l>
    <l>mizu no oto</l>
  </lg>
```

[63] At the time of writing, this W3C module has the status of a candidate recommendation: see further ?

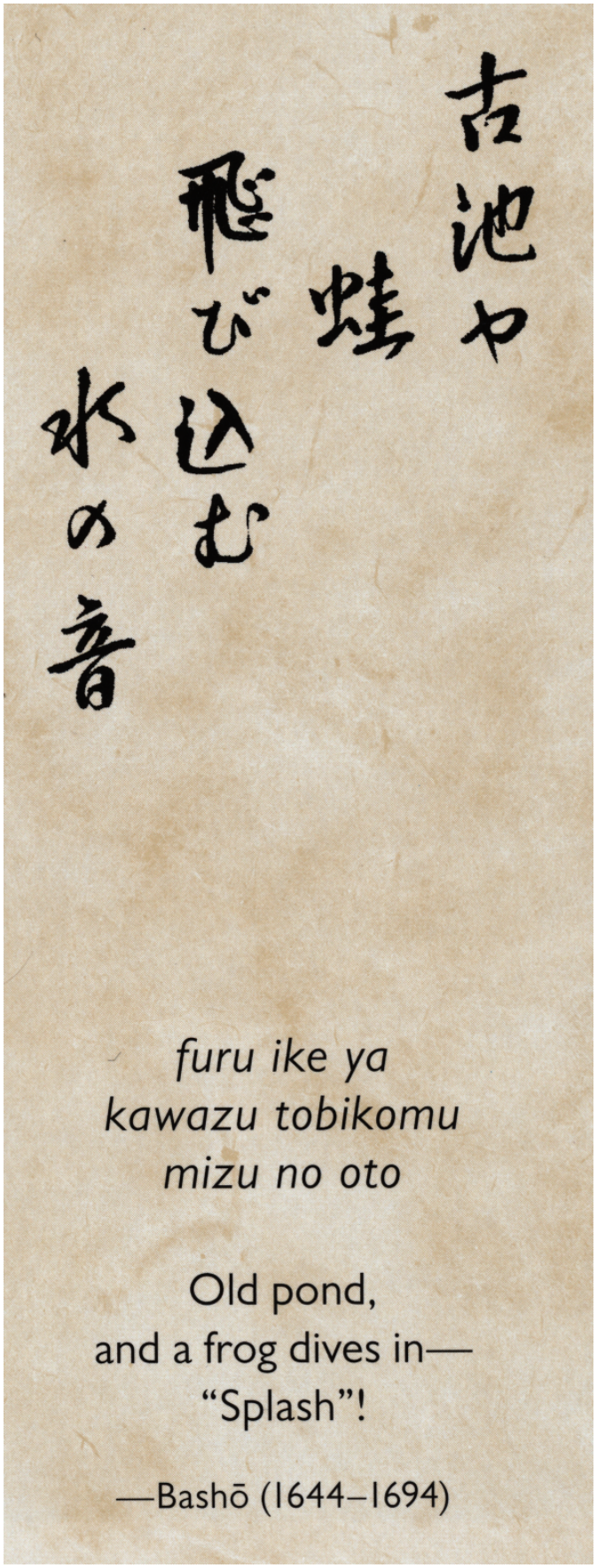



Figure 6: Taken from p.42 of *Haiku: Japanese Art and Poetry*. Judith Patt, Michiko Warkentyne (calligraphy) and Barry Till. 2010.

```
  <lg  xml:lang="en">
    <l>Old pond,</l>
    <l>and a frog dives in—</l>
    <l>"Splash"!</l>
  </lg>
</div>
```

For the sake of simplicity, we have not attempted to capture in this encoding such aspects as the indenting of lines in the first Japanese version, or the central alignment of the other two versions, nor any other renditional features such as font weight or size etc. The Japanese transcription has `writing-mode: vertical-rl`, which is required because Japanese may be written either in this mode or horizontally. The transcription in romaji uses the attribute *xml:lang* to supply a value of ja-Latn, indicating Japanese written in Latin script. Its *style* attribute specifies a horizontal writing mode; this may seem superfluous, but vertically-written romaji is not unknown.

### 5.7.2 Vertical Text with Embedded Horizontal Text

When Japanese is written vertically, the glyph orientation remains the same as when it is written horizontally. In other words, glyphs are not rotated (although as noted above some different glyphs may be used for some characters, in particular for punctuation which needs to be positioned differently in vertical and in horizontal text). However, it is very common for languages written vertically to have embedded runs of text from languages which are normally written horizontally. This raises the issue of the orientation of the glyphs from the horizontal language. Are they written upright, as they would normally appear in horizontal text runs, or are they rotated? Consider this fragment from a Japanese article about the Indonesian language, which takes the form of a glossary list:

The text-orientation property allows us to indicate whether or not glyphs are rotated. In the following example, we have indicated that the list uses a `vertical-rl` writing mode, but that the orientation of individual glyphs may vary:

```
<list  type="gloss"  xml:lang="ja"
  style="writing-mode: vertical-rl; text-orientation: mixed">
  <label  xml:lang="id">hampir</label>
  <item>「近い、ほとんど」</item>
  <label  xml:lang="id">baru</label>
  <item>「新しい、ばかい」</item>
<!-- … -->
</list>
```

The rule `text-orientation: mixed` specifies that ‘characters from horizontal-only scripts are set sideways, i.e. 90° clockwise from their standard orientation in horizontal text. Characters from vertical scripts are set with their intrinsic orientation’ (fantasai 2014). Since the default value for `text-orientation` is `mixed`, this rule is not strictly required. However, if the Indonesian glyphs (which are roman characters) had been set vertically, like this:

then an encoding like the following could be used to make this explicit:

```
<list  type="gloss"  xml:lang="ja"
  style="writing-mode: vertical-rl; text-orientation: upright">
  <label  xml:lang="id">hampir</label>
  <item>「近い、ほとんど」</item>
  <label  xml:lang="id">baru</label>
  <item>「新しい、ばかい」</item>
<!-- … -->
</list>
```

The rule `text-orientation: upright` specifies that ‘characters from horizontal-only scripts are rendered upright, i.e. in their standard horizontal orientation. Characters from vertical scripts are set with their intrinsic orientation and shaped normally’ (fantasai 2014).

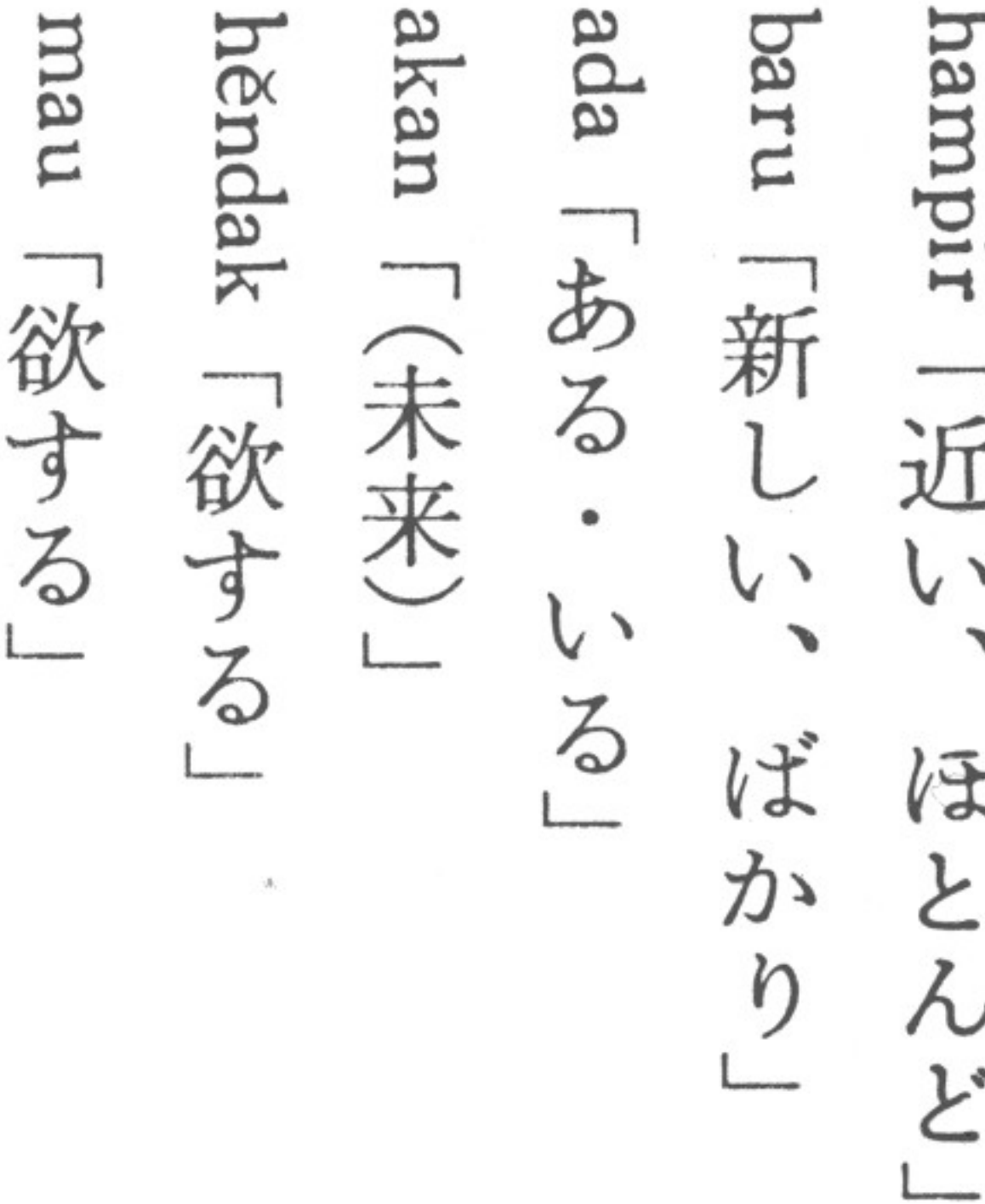



Figure 7: Detail from p.62 of インドネシア語". 崎山理. 1985. 外国語との対照 II. 講座日本語学 11.

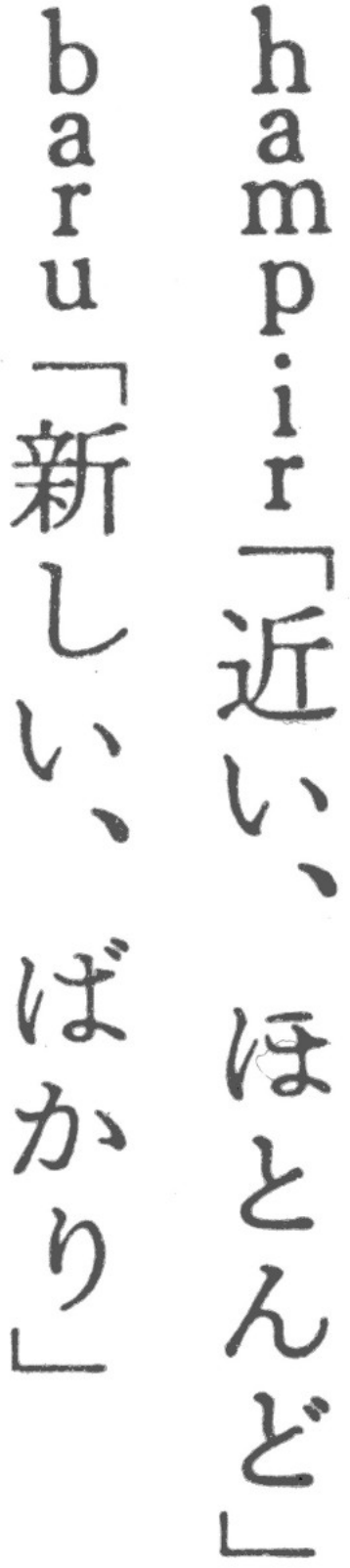



Figure 8: Fragment of previous image with Indonesian glyphs upright.

### 5.7.3 Vertical Orientation in Horizontal Scripts

It is not unusual to see text from horizontal languages written vertically even where no vertically-written script is involved. This example is a fragment from a table of information about agricultural development on Vancouver Island, written in 1855:

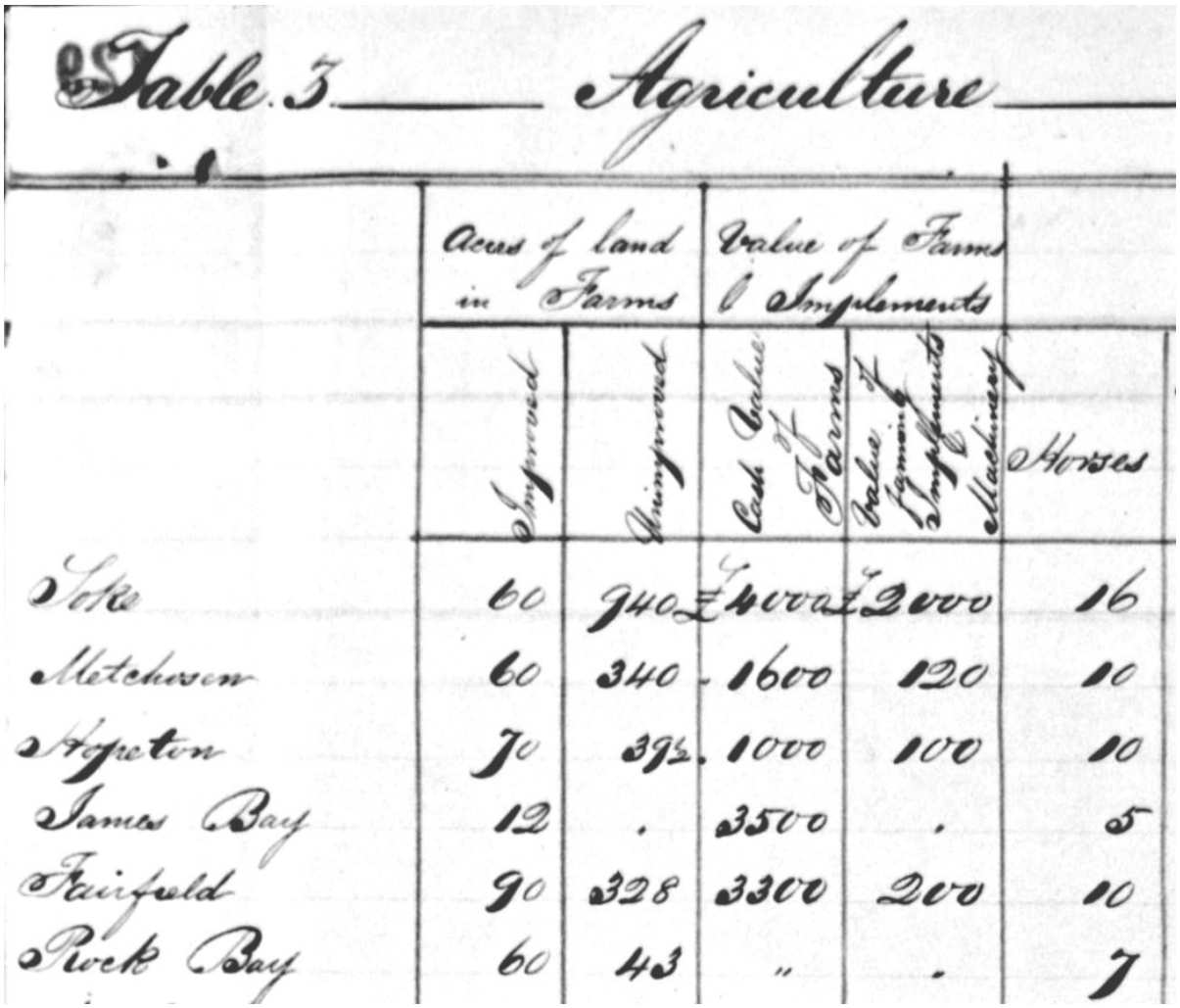


Table. 3 Agriculture

| | Acres of land in Farms | | Value of Farms & Implements | | |
|---|---|---|---|---|---|
| | Improved | Unimproved | Cash Value of Farms | Value of Farming Implements & Machinery | Horses |
| Sooke | 60 | 340 | £4000 | £2000 | 16 |
| Metchosen | 60 | 340 | 1600 | 120 | 10 |
| Hopeton | 70 | 39½ | 1000 | 100 | 10 |
| James Bay | 12 | . | 3500 | . | 5 |
| Fairfield | 90 | 328 | 3300 | 200 | 10 |
| Rock Bay | 60 | 43 | " | . | 7 |

Figure 9: Enclosure with *Despatch to London* 10048, CO 305/6, p. 131v from `https://bcgenesis.uvic.ca/V55116.html`

Four of the subheading cells in this fragment contain English text written vertically, bottom-to-top, to conserve space on the page. To describe this sort of phenomenon, we can use the `text-orientation` property again:

```
text-orientation: mixed | upright | sideways-right | sideways-left |
sideways | use-glyph-orientation
```

For full details on this property, we refer the reader to the CSS Writing Modes specification. For the present example, we will make use only of the 'sideways-left' value, which 'causes text to be set as if in a horizontal layout, but rotated 90° counter-clockwise.' We might encode the third of the four cells containing vertical text like this:

```
<cell  style="writing-mode: vertical-lr; text-orientation: sideways-left">
  <lb/>Cash Value
<lb/>of
<lb/>Farms
```

```
</cell>
```

The `writing-mode` property captures the fact that the script is written vertically, and its lines are to be read from left to right (so the line containing 'of' is to the right of that containing 'Cash value'), while the `text-orientation` value encodes the orientation (rotated 90° counter-clockwise). We might also add `text-align: center` to the style, to express the fact that the text is centrally-aligned.

### 5.7.4 Bottom-to-top Writing

Of the rather small number of scripts which appear to be written bottom-to-top, perhaps the best-known is Ogham, an alphabet used mainly to write Archaic Irish. Ogham is typically found inscribed along the edge of a standing stone, starting at its base. The CSS Writing Modes specification does not explicitly distinguish between vertical scripts which are written top-to-bottom and those which are written bottom-to-top. Instead, such bottom-to-top scripts are best treated as left-to-right horizontal scripts, oriented vertically because of the constraints of the medium on which they are inscribed. Such scripts are analogous to the vertical English text-runs in the table cells in the example above, and can be handled in exactly the same manner (`writing-mode: vertical-lr; text-orientation: sideways-left`). In cases where writing follows a curved path (such as Ogham running around the edge of a stone), a meticulous encoder might resort to the use of SVG to describe the path, rather than treating the phenomenon as a writing mode.

### 5.7.5 Mixed Horizontal Directionality

Returning to our previous simple example

```
  The Arabic term قلم رصاص means "pencil".
```

we could use the direction property to make directionality explicit:
`direction: ltr | rtl`

```
<s  xml:lang="en"  style="direction: ltr">The Arabic term
<term  xml:lang="ar"
   style="direction: rtl; unicode-bidi: embed">قلم رصاص</term> means "pencil".</s>
```

The use of the `direction` property to record the observed directionality of the text is unambiguous, even though it is (as we noted above) superfluous. The use of the `unicode-bidi` property here may require some explanation. By default this property has the value 'normal', the effect of which in this context would be to ignore any value supplied for the direction property. The CSS Writing Modes specification stipulates that the direction property 'has no effect on bidi reordering when specified on inline boxes whose `unicode-bidi` property's value is 'normal', because the element does not open an additional level of embedding with respect to the bidirectional algorithm.'

Mixed horizontal directionality is very common in languages such as Arabic and Hebrew, particularly when numbers (which are always given LTR) or phrases from LTR languages are embedded. It is not impossible, though quite unusual, for ambiguities to arise in such situations, which may give rise to the parts of a document being displayed in unexpected ways that do not correspond to the natural reading order. A more detailed discussion of this issue from an HTML perspective is provided by a W3C Internationalization Working Group report Inline markup and bidirectional text in HTML.

### 5.7.6 Summary

For most texts, information about text directionality need not be explicitly encoded in a TEI text, either because it follows unambiguously from *xml:lang* values, or because it can be expected to be handled unequivocally by the Unicode Bidi Algorithm. Where it is considered important to encode such information, properties and values taken from the CSS Writing Modes module may be used by means of the global TEI *style* attribute (or using the TEI <rendition> element, linked with the *rendition* attribute). Most phenomena can be well described in this way; of those which cannot, other approaches based on the CSS Transforms module are presented in the next section.

## 5.8 Text Rotation

In what follows, we examine a range of textual phenomena which in some ways appear very similar to those examined above, and even overlap with them. We can categorize these as text transformation features, and suggest some strategies for encoding them based on the properties detailed in the CSS Transforms (Fraser et al 2013) specification. This CSS module provides a complex array of properties, values and functions which can be used to rotate, skew, translate and otherwise transform textual and graphical objects. We can borrow this vocabulary in order to describe textual phenomena in a precise manner.

We begin with a simple example of a rotational transform:

Here a block of text has been rotated around its z-axis. This is clearly not a 'writing mode'; the writing mode for this text is horizontal, left to right. Furthermore, even if we wished to treat this as a writing mode, we could not do so, because there is no way to use writing modes properties to describe an text orientation which is angled at 45 degrees; no human languages are consistently written in this orientation. It is more appropriate to treat this as a rotational transformation. We can do this using two properties: `transform` and `transform-origin`. (Both of these properties have quite complex value sets, and we will not look at all of them here. See the specification for full details.)

The `transform` property takes as its value one or more of the transform functions, one of which is the function `rotateZ()`:

```
<ab  style="transform:rotateZ(-45deg)">TEI-C.ORG</ab>
```

Any rotation must take place clockwise around an axis positioned relative to the element being rotated, and the `transform-origin` property can be used to specify the pivot point. By default, the value of `transform-origin` is '50% 50%', the point at the centre of the element, but these values can be changed to reflect rotation around a different origin point. (The TEI <zone> element also bears an attribute *rotate* which can specify rotation in degrees around the z-axis, but it is not available for any other element.)

A block of text may also be rotated about either of its other axes. For example, this shows rotation around the Y (vertical) axis:

```
<ab  style="transform:rotateY(45deg)">TEI-C.ORG</ab>
```

These are obviously trivial examples, but similar features do appear in historical texts. George Herbert's *The Temple* includes two stanzas headed 'Easter Wings' which are both normally printed in a rotated form so that they represent a pair of wings:

This could be encoded thus:

```
<lg  style="transform:rotateZ(90deg)">
  <l>My tender age in ſorrow did beginne:</l>
  <l>And ſtill with ſickneſſes and ſhame</l>
<!-- ... -->
</lg>
```

We might also argue that this is in fact a vertical writing mode by supplying `writing-mode: vertical-rl; text-orientation: sideways-right` as the value for the *style* attribute in the preceding example.

Rotation is also useful as a method of handling a true writing mode which is not covered by the CSS Writing Modes: *boustrophedon*. This is a writing mode common in inscriptions in Latin, Greek and other languages, in which alternate lines run from left to right and from right to left[64]. Right-to-left lines in boustrophedon have another unexpected feature: their glyphs are reversed, so that these lines appear as 'mirror writing', as in the following ancient Greek inscription:

This might be transcribed as follows (ignoring word boundaries for the moment):

```
<ab>
  <lb/>HEPMONTINA
<lb/>
  <seg  style="rotateY(180deg)">KA0EONΠOT0EM</seg>
  <lb/>ENOΣYENEAF
<lb/>
  <seg  style="rotateY(180deg)">OIYENOITIEKK</seg>
  <lb/>PETAIAΣONA
<lb/>
  <seg  style="rotateY(180deg)">ΣIMOΣOTTAIE</seg>
  <lb/>AΣΣAI

</ab>
```

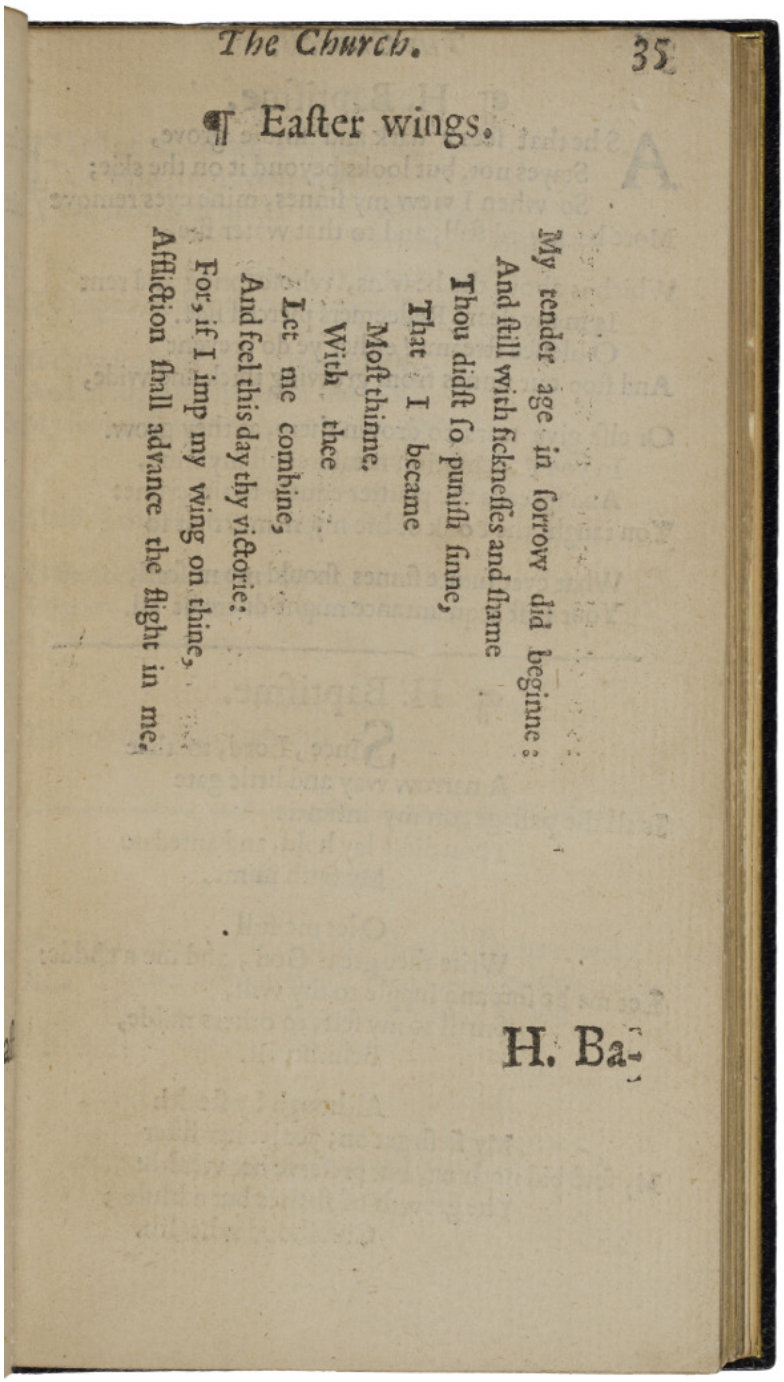
The 180-degree rotation around the Y (vertical) axis here describes what is happening in the RTL lines in boustrophedon; the order of glyphs is reversed, and so is their individual orientation (in fact, we see them 'from the back', as it were). <seg> elements have been used here because these are clearly not 'lines' in the sense of poetic lines; the text is continuous prose, and the division into separate lines is incidental.

There are obviously some unsatisfactory aspects of this manner of encoding boustrophedon. In the inscription above, some words are split across two lines, so if we wished to tag both words and the right-to-left phenomena, one hierarchy would have to be privileged over the other. By using a transform function rather than a writing mode property, we are apparently suggesting that boustrophedon is not in fact a writing mode, whereas it clearly is. But the CSS Writing Modes specification does not provide support for boustrophedon, because it is a rather obscure historical phenomenon; using a rotational transform is one practical alternative.

## 5.9 Caveat

As with other parts of the CSS specification, the intended effect of CSS Transforms properties and values is defined with reference to a specific Visual formatting model; the language is designed to describe how an HTML document

[64] The name is taken from the Greek βουστροφηδόν, meaning 'ox-turning' from βοῦς (an ox) and στροφή ('turn'); that is, turning as an ox does when pulling a plough.



The Church. 35

¶ Easter wings.

My tender age in sorrow did beginne:
And still with sicknesses and shame
Thou didst so punish sinne,
That I became
Most thinne.
With thee
Let me combine,
And feel this day thy victorie:
For, if I imp my wing on thine,
Affliction shall advance the flight in me.

H. Ba

Figure 10: Page 35 of George Herbert's *The Temple* (1633), from a copy in the Folger Library.

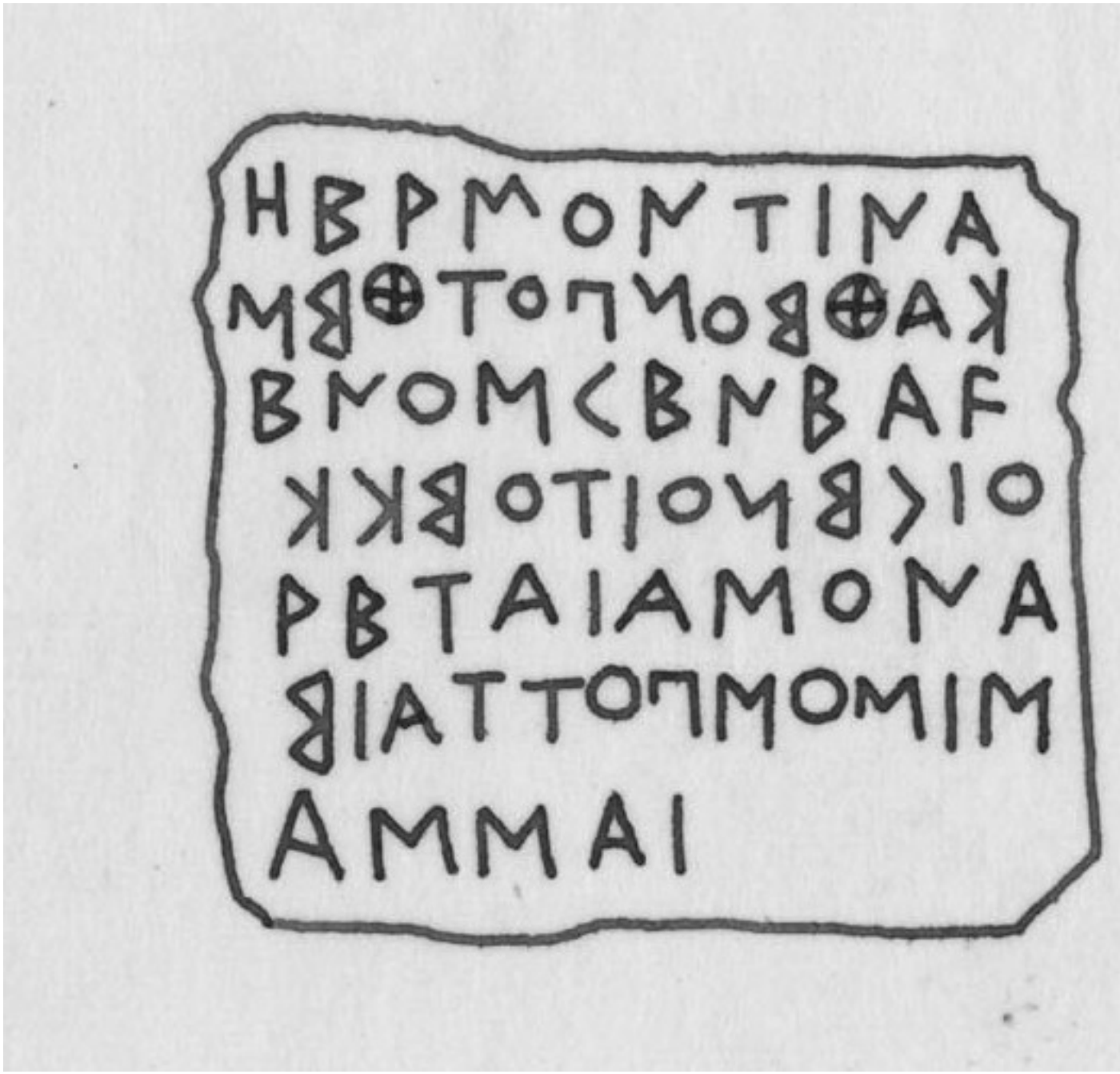


Figure 11: Leaden plaque bearing an inquiry by Hermon from the oracular precinct at Dodona. (L.H. Jeffery Archive)

should be formatted. This is not, of course, the case for the TEI, which lacks any explicit processing or formatting model, and attempts to define objects as far as possible without consideration of their visual appearance. As long as the properties and values from the CSS Transforms module are used as a convenient, well-specified descriptive language to capture features of a text, without any expectation of using them directly and reliably for rendering, this is not particularly problematic. CSS provides a useful and well-defined vocabulary to describe many aspects of the appearance of source texts, benefitting particularly from the clarity of definition provided by the specification. However, if there is any expectation of using this information to render a text in a predictable and accurate way, it will be essential to provide enough styling information throughout the document hierarchy to resolve all ambiguities with regard to size, positioning, block status, etc. before any element undergoes a transform operation.

## 5.10 Formal Definition

The gaiji module described in this chapter makes available the following components:

**Module gaiji: Character and glyph documentation**

- *Elements defined*: char charDecl g glyph localProp mapping unicodeProp unihanProp
- *Classes defined*: att.gaijiProp

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 6 Verse

This module is intended for use when encoding texts which are entirely or predominantly in verse, and for which the elements for encoding verse structure already provided by the core module are inadequate.

The tags described in section *3.13.1. Core Tags for Verse* include elements for the encoding of verse lines and line groups such as stanzas: these are available for any TEI document, irrespective of the module it uses. Like the modules for prose and for drama, the module for verse additionally makes use of the module defined in chapter *4. Default Text Structure* to define the basic formal structure of a text, in terms of <front>, <body> and <back> elements and the text-division elements into which these may be subdivided.

The module for verse extends the facilities provided by these modules in the following ways:

- a special purpose <caesura> element is provided, to allow for segmentation of the verse line (see section *6.2. Components of the Verse Line*)
- a set of attributes is provided for the encoding of rhyme scheme and metrical information (see sections *6.4. Rhyme and Metrical Analysis* and *6.5. Rhyme*)
- a special purpose <rhyme> element is provided to support simple analysis of rhyming words (see section *6.5. Rhyme*)

## 6.1 Structural Divisions of Verse Texts

Like other kinds of text, texts written in verse may be of widely differing lengths and structures. A complete poem, no matter how short, may be treated as a free-standing text, and encoded in the same way as a distinct prose text. A group of poems functioning as a single unit may be encoded either as a <group> or as a <text>, depending on the encoder's view of the text. For further discussion, including an example encoding for a verse anthology, see chapter *4. Default Text Structure*.

Many poems consist only of ungrouped lines. This short poem by Emily Dickinson is a simple case:

```
<text>
  <front>
    <head>1755</head>
  </front>
  <body>
    <l>To make a prairie it takes a clover and one bee,</l>
    <l>One clover, and a bee,</l>
    <l>And revery.</l>
    <l>The revery alone will do,</l>
    <l>If bees are few.</l>
  </body>
</text>
```

Often, however, lines are grouped, formally or informally, into stanzas, verse paragraphs, etc. The <lg> element defined in the core tag set (in section *3.13.1. Core Tags for Verse*) may be used for all such groupings. It may thus serve for informal groupings of lines such as those of the following example from Allen Ginsberg:

```
<text>
  <body>
    <head>My Alba</head>
    <lg>
      <l>Now that I've wasted</l>
      <l>five years in Manhattan</l>
      <l>life decaying</l>
      <l>talent a blank</l>
    </lg>
    <lg>
      <l>talking disconnected</l>
      <l>patient and mental</l>
```

```
     <l>sliderule and number</l>
     <l>machine on a desk</l>
   </lg>
  </body>
</text>
```

It may also be used to mark the verse paragraphs into which longer poems are often divided, as in the following example from Samuel Taylor Coleridge's *Frost at Midnight*:

```
<lg>
  <l>The Frost performs its secret ministry,</l>
  <l>Unhelped by any wind. ...</l>
  <l>Whose puny flaps and freaks the idling Spirit</l>
  <l>By its own moods interprets, every where</l>
  <l>Echo or mirror seeking of itself,</l>
  <l  part="I">And makes a toy of Thought.</l>
</lg>
<lg>
  <l  part="F">But O! how oft,</l>
  <l>How oft, at school, with most believing mind</l>
  <l>Presageful, have I gazed upon the bars,</l>
  <l>To watch that fluttering <hi>stranger</hi>! ... </l>
</lg>
<lg>
  <l>Dear Babe, that sleepest cradled by my side,</l>
</lg>
```

Note, in the above example, the use of the *part* attribute on the <l> element, where a verse line is broken between two line groups, as discussed in section *3.13.1. Core Tags for Verse*.

Most typically, however, the <lg> element is used to mark the highly regular line groups which characterize stanzaic and similar verse forms, as in the following example from Chaucer:

```
<lg>
  <l>Sire Thopas was a doghty swayn;</l>
  <l>White was his face as payndemayn,</l>
  <l>His lippes rede as rose;</l>
  <l>His rode is lyk scarlet in grayn,</l>
  <l>And I yow telle in good certayn,</l>
  <l>He hadde a semely nose.</l>
</lg>
<lg>
  <l>His heer, his ber was lyk saffroun,</l>
  <l>That to his girdel raughte adoun;</l>
</lg>
```

Like other text-division elements, <lg> elements may be nested hierarchically. For example, one particularly common English stanzaic form consists of a quatrain or sestet followed by a couplet. The <lg> element may be used to encode both the stanza and its components, as in the following example from Byron:

```
<lg  type="stanza">
  <lg  type="sestet">
    <l>In the first year of Freedom's second dawn</l>
```

```
    <l>Died George the Third; although no tyrant, one</l>
    <l>Who shielded tyrants, till each sense withdrawn</l>
    <l>Left him nor mental nor external sun:</l>
    <l>A better farmer ne'er brushed dew from lawn,</l>
    <l>A worse king never left a realm undone!</l>
  </lg>
  <lg  type="couplet">
    <l>He died — but left his subjects still behind,</l>
    <l>One half as mad — and t'other no less blind.</l>
  </lg>
</lg>
```

Note the use of the *type* attribute to name the type of unit encoded by the <lg> element; this attribute is common to all members of the `att.divLike` class (see section *4.1.1. Un-numbered Divisions*).[65] When used on <lg>, the *type* attribute is intended solely for conventional names of different classes of text block. For systematic analysis of metrical and rhyme schemes, use the *met* and *rhyme* attributes, for which see below, section *6.4. Rhyme and Metrical Analysis*.

As a further example, consider the Shakespearean sonnet. This may be divided into two parts: a concluding couplet, and a body of twelve lines, itself subdivided into three quatrains:

```
<text>
  <body>
    <lg>
      <lg  type="quatrain">
        <l>My Mistres eyes are nothing like the Sunne,</l>
        <l>Currall is farre more red, then her lips red</l>
        <l>If snow be white, why then her brests are dun:</l>
        <l>If haires be wiers, black wiers grown on her head:</l>
      </lg>
      <lg  type="quatrain">
        <l>I have seene Roses damaskt, red and white,</l>
        <l>But no such Roses see I in her cheekes,</l>
        <l>And in some perfumes is there more delight,</l>
        <l>Then in the breath that from my Mistres reekes.</l>
      </lg>
      <lg  type="quatrain">
        <l>I love to heare her speake, yet well I know,</l>
        <l>That Musicke hath a farre more pleasing sound:</l>
        <l>I graunt I never saw a goddesse goe,</l>
        <l>My Mistres when shee walkes treads on the ground.</l>
      </lg>
    </lg>
    <lg  type="couplet">
      <l>And yet by heaven I think my love as rare,</l>
      <l>As any she beli'd with false compare.</l>
    </lg>
  </body>
</text>
```

Particularly lengthy poetic texts are often subdivided into units larger than stanzas or paragraphs, which may themselves be subdivided. Spenser's *Faery Queene*, for example, consists of twelve 'books' each of which contains a prologue followed by twelve 'cantos'. Each prologue and each canto consists of nine-line 'stanzas', each of which follows the same regular pattern. Other examples in the same tradition are easy to find.

[65] For discussion of other attributes of this class, see *4.1.4. Partial and Composite Divisions*.

Large structures of this kind are most conveniently represented by  or <div1> elements, as described in section *4.1. Divisions of the Body*. Thus the start of the *Faerie Queene* might be encoded as follows:

```
<body>
  <div  n="I"  type="book">
    <div  n="1"  type="canto">
      <lg  n="I.1.1"  type="stanza">
        <l>A Gentle Knight was pricking on the plain</l>
        <l>Y cladd in mightie armes and silver shielde,</l>
      </lg>
    </div>
  </div>
</body>
```

The encoder must choose at which point in the hierarchy of structural units to introduce <lg> elements rather than a yet smaller  element: it would (for example) also be possible to encode the above example as follows:

```
<body>
  <div  n="I"  type="book">
    <div  n="I.1"  type="canto">
      <div  n="I.1.1"  type="stanza">
        <l>A gentle knight was pricking on the plain</l>
        <l>Ycladd in mightie armes and silver shielde,</l>
      </div>
    </div>
  </div>
</body>
```

One reason for using  rather than <lg> elements is that the former may contain non-metrical elements, such as epigraphs or dedications and other members of the `model.divTop` class, whereas <lg> elements may contain only headings or metrical lines.

## 6.2 Components of the Verse Line

It is often convenient for various kinds of analysis to encode subdivisions of verse lines. The general purpose <seg> element defined in the tag set for segmentation and alignment (section *17.3. Blocks, Segments, and Anchors*) is provided for this purpose:

`<seg>` (arbitrary segment) represents any segmentation of text below the 'chunk' level.

To use this element together with the module for verse, the module for segmentation and alignment must also be enabled as further described in section *1.2. Defining a TEI Schema*.

In Old and Middle English alliterative verse, individual verse lines are typically split into half lines. The <seg> element may be used to mark these explicitly, as in the following example from Langland's *Piers Plowman*:

```
<l>
  <seg>In a somer
     seson,</seg>
  <seg>whan softe was the sonne,</seg>
</l>
<l>
  <seg>I shoop me into shroudes</seg>
  <seg>as I a sheep were,</seg>
</l>
<l>
  <seg>In habite as an heremite </seg>
```

```
  <seg>unholy of werkes,</seg>
</l>
<l>
  <seg>Went wide in this world </seg>
  <seg>wondres to here.</seg>
</l>
```

The <seg> element can be nested hierarchically, in the same way as the <lg> element, down to whatever level of detailed structure is required. In the following example, the line has been divided into *feet*, each of which has been further subdivided into syllables.[66]

```
<l>
  <seg  type="foot">
    <seg  type="syll">Ar</seg>
    <seg  type="syll">ma </seg>
    <seg  type="syll">vi</seg>
  </seg>
  <seg  type="foot">
    <seg  type="syll">rum</seg>
    <seg  type="syll">que </seg>
    <seg  type="syll">ca</seg>
  </seg>
  <seg  type="foot">
    <seg  type="syll">no </seg>
    <seg  type="syll">Tro</seg>
  </seg>
  <seg  type="foot">
    <seg  type="syll">iae </seg>
    <seg  type="syll">qui </seg>
  </seg>
  <seg  type="foot">
    <seg  type="syll">pri</seg>
    <seg  type="syll">mus </seg>
    <seg  type="syll">ab </seg>
  </seg>
  <seg  type="foot">
    <seg  type="syll">or</seg>
    <seg  type="syll">is </seg>
  </seg>
</l>
```

The <seg> element may be used to identify any subcomponent of a line which has content; its type attribute may characterize such units in any way appropriate to the needs of the encoder. For the specific case of labeling each foot with its formal type ('dactyl', 'spondee', etc.), and each syllable with its metrical or prosodic status (syllables bearing primary or secondary stress, long syllables, short syllables), however, the specialized attributes *met* and *real* are defined, which provide a more systematic framework than the *type* attribute; see section *6.4. Rhyme and Metrical Analysis* below.

In classical verse, a hexameter like that above may also be formally divided into two *cola* or 'hemistiches'. This example provides a typical case, in that the boundary of the first colon falls in the middle of one of the feet (between the syllables 'no' and 'Tro'). If both kinds of segmentation are required, the *part* attribute might be used to mark the overlapping structure as follows.

[66] As elsewhere in these Guidelines, this example has been formatted for clarity of exposition rather than correct display. Note in particular that whether an XML processor retains whitespace within the <seg> element or not (this can be configured by means of the *xml:space* attribute) this example will still require additional processing, since whitespace should be retained for the lower level <seg> elements (those of type syll) but not for the higher level one (those of type foot).

```
<l>
  <seg type="hemistich">
    <seg type="foot">
      <seg type="syll">Ar</seg>
      <seg type="syll">ma </seg>
      <seg type="syll">vi</seg>
    </seg>
    <seg type="foot">
      <seg type="syll">rum</seg>
      <seg type="syll">que </seg>
      <seg type="syll">ca</seg>
    </seg>
    <seg type="foot" part="I">
      <seg type="syll">no </seg>
    </seg>
  </seg>
  <seg type="hemistich">
    <seg type="foot" part="F">
      <seg type="syll">Tro</seg>
    </seg>
    <seg type="foot">
      <seg type="syll">iae </seg>
      <seg type="syll">qui </seg>
    </seg>
  </seg>
</l>
```

Instead of using the *part* attribute on the <seg> element, it might be simpler just to mark the point at which the caesura occurs. An additional element is provided for analyses of this kind, in which what is to be marked are points 'between the words', which have some significance within a verse line:

`<caesura>` marks the point at which a metrical line may be divided.

In classical prosody, the *caesura*, which occurs within a foot, is distinguished from a *diaeresis*, which occurs on a foot boundary (not to be confused with the division of a diphthong into two syllables, or the diacritic symbol used to indicate such division, each of which is also termed *diaeresis*). This distinction is rarely made nowadays, the term *caesura* being used for any division irrespective of foot boundaries. No special-purpose `<diaeresis>` element is therefore provided.

As an example of the <caesura> element, we refer again to the example from Langland. An encoder might choose simply to record the location of the caesura within each line, rather than encoding each half-line as a segment in its own right, as follows:

```
<l>In a somer seson,
<caesura/> whan softe was the sonne, </l>
<l>I shoop me into shroudes <caesura/> as I a sheep were, </l>
<l>In habite as an heremite <caesura/> unholy of werkes, </l>
<l>Went wide in this world <caesura/> wondres to here. </l>
```

Logically, the opposite of caesura might be considered to be *enjambement*. When the `verse` module is included in a schema, an additional class called `att.enjamb` is defined as follows:

`att.enjamb` (enjambement) provides attributes that may be used to indicate enjambement of the parent element.

`@enjamb` (enjambement) indicates that the end of a verse line is marked by enjambement.

The following lines demonstrate the use of the *enjamb* attribute to mark places where there is a discrepancy between the boundaries of the <l> elements and the syntactic structure of the verse (a discrepancy of some significance in some schools of verse):

```
<l  enjamb="y">Un astrologue, un jour, se laissa choir</l>
<l>Au fond d'un puits.</l>
```

## 6.3 Encoding Textual Structures Across Verses

It is possible that certain textual structures may span multiple lines of verse, either by incorporating more than one, or by crossing line hierarchy. This is common, for example, when lines contain reported thought or speech (i.e. <said>), or other forms of quotation (i.e. <q>). For these cases, it is recommended practice to fragment and reconstruct the elements representing the textual structures.

The following example from Margaret Cavendish's *Nature's Pictures* shows speech encoded across two lines reconstructed by chaining elements with *prev* and *next* attributes:

```
<lg  type="couplet">
  <l>
    <said  xml:id="eg1-said1"
      next="#eg1-said2">Our lives</said>, ſaid he,
  <said  xml:id="eg1-said2"
      next="#eg1-said3"  prev="#eg1-said1">wee'll give before we yield</said>,
  </l>
  <l>
    <said  xml:id="eg1-said3"
      prev="#eg1-said2">Wee'll win your battles, or dye in the field</said>.
  </l>
</lg>
```

Alternatively, the elements may be reconstructed with stand-off markup using the element <join>:

```
<lg  type="couplet">
  <l>
    <seg  xml:id="eg2-said1">Our lives</seg>, ſaid he,
  <seg  xml:id="eg2-said2">wee'll give before we yield</seg>,
  </l>
  <l>
    <seg  xml:id="eg2-said3">Wee'll win your battles, or dye in the field</seg>.
  </l>
</lg>
<!-- Elsewhere in the document -->
<p>
  <join  result="said"  scope="root"
    target="#eg2-said1 #eg2-said2 #eg2-said3"/>
</p>
```

A more general discussion of these and other strategies to deal with fragmentation and reconstruction appears in section *21.3. Fragmentation and Reconstitution of Virtual Elements*.

## 6.4 Rhyme and Metrical Analysis

When the module for verse is in use, the following additional attributes are available to record information about rhyme and metrical form:

`att.metrical` defines a set of attributes that certain elements may use to represent metrical information.

`@met` (metrical structure, conventional) contains a user-specified encoding for the conventional metrical structure of the element.

`@real` (metrical structure, realized) contains a user-specified encoding for the actual realization of the conventional metrical structure applicable to the element.

`@rhyme` (rhyme scheme) specifies the rhyme scheme applicable to a group of verse lines.

These attributes may be attached to the <lg> element, or to the higher-level text-division elements , <div1>, etc. In general, the attributes should be specified at the highest level possible; they may not however be specifiable at the highest level if some of the subdivisions of a text are in prose and others in verse. All these attributes may also be attached to the <l> and <seg> elements, but the default notation for the *rhyme* attribute has no defined meaning when specified on <l> or <seg>. The value for these attributes may take any form desired by the encoder, but the nature of the notation used will determine how well the attribute values can be processed by automatic means.

The primary function of the metrical attributes is to encode the conventional metrical or rhyming structure within which the poet is working, rather than the actual prosodic realization of each line; the latter can be recorded using the *real* attribute, as further discussed below. A simple mechanism is also provided for recording the actual realization of a rhyme pattern; see *6.5. Rhyme*.

### 6.4.1 Sample Metrical Analyses

As a simple example of the use of these attributes, consider the following lines from Pope's 'Essay on Criticism':

```
<div type="book" n="1" met="-+|-+|-+|-+|-+/"
  rhyme="aa">
  <lg n="1" type="paragraph">
    <l>'Tis hard to say, if greater Want of Skill</l>
    <l>Appear in <hi>Writing</hi> or in <hi>Judging</hi> ill;</l>
    <l>But, of the two, less dang'rous is th'Offence,</l>
    <l>To tire our <hi>Patience</hi>, than mis-lead our <hi>Sense</hi>:</l>
  </lg>
</div>
```

This text is written entirely in *heroic couplets*; each line is an iambic pentameter (which, using a common notation, can be described with the formula -+|-+|-+|-+|-+/, each - denoting a metrically unstressed syllable, each + a metrically stressed one, each | a foot boundary, and the / a line-end), and the couplets rhyme (which can be represented with the conventional formula aa).

Because both rhyme pattern and metrical form are consistent throughout the poem, they may be conveniently specified on the  element; the values given for the attributes will be inherited by any metrical unit contained within the  elements of this poem, and must be interpreted in the appropriate way.

Since the notation used in the *met*, *real*, and *rhyme* attributes is user-defined, no binding description can be given of its details or of how its interpretation must proceed. (A default notation is provided for the *rhyme* attribute, which however the encoder can replace with another; see section *6.5. Rhyme*.) It is expected, however, that software should be able to support these attributes in useful ways; the more intelligent the software is, and the more knowledge of metrics is built into it, the better it will be able to support these attributes. In the extract given above, for example, the *met* and *rhyme* attribute values specified on the  element are inherited directly by the <lg> elements nested within it. Since the *met* value specifies the metrical form of a single verse line, the structure of the <lg> as a whole is understood to involve as many repetitions of the pattern as there are lines in the verse paragraph. The same attribute value, when inherited in turn by the <l> element, must be understood *not* to repeat. With sufficiently sophisticated software, segments within the line might even be understood as inheriting precisely that portion of the formula which applies to the segment in question; this will, however, be easier to accomplish for some languages than for others.

The *rhyme* attribute in this example uses the default notation to specify a rhyme scheme applicable only to pairs of lines. As elsewhere, the default notation for the *rhyme* attribute has no meaning for metrical units at the line level or below. In verse forms where line-internal rhyme is structurally significant, e.g. in some skaldic poetry, the default notation is incapable of expressing the required information, since the rhyme pattern may need to be specified for units smaller than the line. In such cases, a user-specified rhyme notation must be substituted for the default notation, or else the rhyme pattern must be described using some alternative method (e.g. by using the <link> mechanism described below).

The precise semantics of the *met* attribute and the inferences which software is expected or able to draw from it, are implementation-dependent; so are the semantics and processing of the *rhyme* attribute, when user-specified notations are used.

A formal definition of the significance of each component of the pattern given as the value of the *met* attribute may be provided in the <metDecl> element within the <encodingDesc> element in the TEI header (see section *6.6. Metrical Notation Declaration*). The encoder is free to invent any notation appropriate to his or her analytic needs, provided that it is adequately documented in this element. The notation may define metrical components using invented or traditional names (such as 'iamb' or 'hexameter') or in terms of basic units such as codes for stressed or unstressed syllables, or a combination of the two.

The *real* (for 'realization') attribute may optionally be specified to indicate any deviation from the pattern defined by the *met* attribute which the encoder wishes to record. By default, the *real* attribute has the same value as the *met* attribute on the same element; it is only necessary to provide an explicit value when the realization differs in some way from the abstract metrical pattern. The tension between conventional metrical pattern and its realization may thus be recorded explicitly. For example, many readers of the above passage would stress the word 'But' at the beginning of the third line rather than the word 'of' following it, as the metrical pattern would normally require. This variation might be encoded as follows:

```
<l  real="+-|-+|-+|-+|-+">But, of the two, ...</l>
```

Where the *real* attribute is used to over-ride the default or conventional metrical pattern, it applies only to the element on which it is specified. The default pattern for any subsequent lines is unaffected.

As it happens, this particular kind of variation is very common in the English iambic pentameter—it even has a name: *trochaic substitution*—an encoder might therefore choose to regard this not as an instance of a variant realization, but as an instance of a variant metrical form:

```
<l  met="+-|-+|-+|-+|-+">But, of the two, ...</l>
```

Alternatively, a different metrical notation might be defined, in which this kind of variation was permitted throughout the text.

In choosing whether to over-ride a metrical specification in this way or by using the *real* attribute, the encoder is required to determine whether the change is a systematic or conventional one (as in this example) or an occasional variation, perhaps for local effect. In the following example, from Goethe's 'Auf dem See', the variation is a matter of local realization:

```
<lg  type="chevy-chase-stanza"
  met="-+-+-+-+/-+-+-+"  rhyme="ababcdcd">
  <l  n="1"> Und frische Nahrung, neues Blut</l>
  <l  n="2"  real="+--+-+"> Saug' ich aus freier Welt;</l>
  <l  n="3"  real="+--+-+-+"> Wie ist Natur so hold und gut,</l>
  <l  n="4"  real="---+-+"> Die mich am Busen hält!</l>
  <l  n="5"> Die Welle wieget unsern Kahn</l>
  <l  n="6"> Im Rudertakt hinauf,</l>
  <l  n="7"> Und Berge, wolkig himmelan,</l>
  <l  n="8"> Begegnen unserm Lauf.</l>
</lg>
```

On the other hand, the famous inserted alexandrine in Pope's 'Essay on Criticism', might be encoded as follows:

```
<l  n="356"> A
needless alexandrine ends the song, </l>
<l  n="357"  met="-+|-+|-+|-+|-+|-+"
```

```
  real="++|-+|-+|+-|++|-+"> That, like a wounded
snake, drags its slow length along. </l>
```

Here the *met* attribute indicates that a different metrical convention (the alexandrine) is in force, while the *real* attribute indicates that there is a variation from that convention. As with many other aspects of metrical analysis, however, this is of necessity an entirely interpretive judgment.

### 6.4.2 Segment-Level versus Line-level Tagging

The examples given so far have encoded information about the realization of metrical conventions at the level of the whole verse-line. This has obvious advantages of simplicity, but the disadvantage that any deviation from metrical convention is not marked at its precise point of occurrence in the text. Greater precision may be achieved, but only at the cost of marking deviant metrical units explicitly. This may be done with the <seg> element, giving the variant realization as the value of the *real* attribute on that element. Using this method, the example given immediately above might be encoded as follows:

```
<l  n="356"> A need<seg  type="foot"  n="2"  real="--">less a</seg>lexandrine ends the
song,</l>
<l  n="357"  met="-+|-+|-+|-+|-+|-+">
  <seg  n="1"  real="++"> That, like </seg> a wounded snake, <seg  n="4"  real="+-">
drags its </seg>
  <seg  n="5"  real="++"> slow length </seg> along.
</l>
```

The marking of the foot boundaries with the symbol | in the *met* attribute value of the <l> element allows the human reader, or a sufficiently intelligent software program, to isolate the correct portion of that attribute value as the default value for the same attribute on the <seg> elements for feet, namely -+. It is of course up to the encoder to decide whether or not to include the *n* attribute of <seg> here, and whether or not also to tag the feet in the line in which there is no deviation from the metrical convention. The ability of software to infer which foot is being marked, if not all are tagged, will depend heavily on the language of the text and the knowledge of prosody built into the software; the fuller and more explicit the markup, the easier it will be for software to handle it. It may prove useful, however, to mark metrical deviations in the manner shown, even if the available software is not sufficiently intelligent to scan lines without aid from the markup. Human readers who are interested in prosody may well be able to exploit the markup in useful ways even with less sophisticated software.

There are circumstances where it may also be useful to use the *met* attribute of <seg>. If we wish to identify the exact location of the different types of foot in the first line of Virgil's *Aeneid*, the text could be encoded as follows (for simplicity's sake the caesura has been omitted):

```
<l>
  <seg  type="foot"  met="+--">Arma vi</seg>
  <seg  met="+--">rumque ca</seg>
  <seg  met="++">no Tro</seg>
  <seg  met="++">iae qui </seg>
  <seg  met="+--">primus ab</seg>
  <seg  met="++"> oris</seg>
</l>
```

An appropriate value of the *met* attribute might also be supplied on the enclosing  element, to indicate that each foot may be made up of a dactyl or a spondee, so that the values given here for *met* at the level of the foot may be considered a series of local variations on this fundamental pattern; in cases like this, of course, the local variations may also be considered aspects of realization rather than of convention, in which case the *real* attribute may be used instead of *met*, if desired.

### 6.4.3 Metrical Analysis of Stanzaic Verse

The method described above may be used to encode quite complex verse forms, for instance various kinds of fixed-form stanzas. Let us take one of Dante's canzoni, in which each stanza except the last has the same combination of eleven-syllable and seven-syllable lines, and the same rhyme scheme:

```
<div type="canzone"
  met="E/E/S/E/S/E/E/S/E/S/E/S/S/E/S/E/E/S/S/E/E" rhyme="abbcdaccbdceeffghhhgg">
  <lg n="1" type="stanza">
    <l n="1">Doglia mi reca nello core ardire</l>
  </lg>
</div>
```

Here the *met* attribute specifies a metrical pattern for each of the twenty-one lines making up a stanza of the *canzone*. Each stanza inherits this definition from the parent  element. The *rhyme* attribute specifies a rhyme scheme for each stanza, in the same way.

In the metrical notation used here, the letter E represents a line containing nine syllables which may or may not be metrically prominent, a tenth which is prominent and an optional non-prominent eleventh syllable. The letter 'S' is used to represent a line containing five syllables which may or may not be metrically prominent, a sixth which is prominent and an optional non-prominent seventh syllable. A suitable definition for this notation might be given by a <metDecl> element like the following:

```
<metDecl type="met" pattern="((E|S)/)+)">
  <metSym value="E" terminal="false">xxxxxxxxx+o</metSym>
  <metSym value="S" terminal="false">xxxxx+o</metSym>
  <metSym value="x">metrically prominent or non-prominent</metSym>
  <metSym value="+">metrically prominent</metSym>
  <metSym value="o">optional non prominent</metSym>
  <metSym value="/">line division</metSym>
</metDecl>
```

As noted above, the metrical pattern specified on the  applies to each <lg> (stanza) element contained within the . In fact however, after seven stanzas of this type, there is a final stanza, known as a *commiato* or envoi, which follows a different metrical and rhyming scheme. The solution to this problem is simply to specify a new *met* attribute on the eighth stanza itself, which will override the default value inherited from parent , as follows:

```
<div met="[...]">
  <lg>
    <l> ... </l>
  </lg>
  <lg type="commiato"
    met="E/S/S/E/S/E/E/S/S/E/E" rhyme="abbccdeeedd">
    <l n="1">Canzone, presso di qui è une donna</l>
  </lg>
</div>
```

Note that, in the same way as for the *real* attribute, over-riding of this kind does not affect subsequent elements at the same hierarchic level. Any <lg> element following the *commiato* above would be assumed to use the same metrical and rhyming scheme as the one preceding the *commiato*. Moreover, although it is quite regular (in the sense that the last stanza of each *canzone* is a *commiato*), the over-riding must be specified for each case.

## 6.5 Rhyme

The *rhyme* attribute is used to specify the rhyme pattern of a verse form. It should not be confused with the <rhyme> element, which is used to mark the actual rhyming word or words:

`<rhyme>` marks the rhyming part of a metrical line.

Like the *met* attribute, the *rhyme* attribute can be used with a user-specified notation documented by the <metDecl> element in the TEI header. Unlike *met*, however, the *rhyme* attribute has a default notation; if this default notation is used, no <metDecl> element need be given.

The default notation for rhyme offers the ability to record patterns of rhyming lines, using the traditional notation in which distinct letters stand for rhyming lines. For a work in rhyming couplets, like the Pope example above, the *rhyme* attribute simply specifies aa, indicating that pairs of adjacent lines rhyme with each other. For a slightly more complex scheme, applicable to groups of four lines, in which lines 1 and 3 rhyme, as do lines 2 and 4, this attribute would have the value abab. The traditional Spenserian stanza has the pattern ababbcbcc, indicating that within each nine line stanza, lines 1 and 3 rhyme with each other, as do lines 2, 4, 5 and 7, and lines 6, 8 and 9.

Non-rhyming lines within such a group may be represented using a hyphen or an x, as in the following example:

```
<lg  rhyme="aa-a">
  <l>Why, all the Saints and Sages who discuss'd</l>
  <l>Of the Two Worlds so learnedly, are thrust</l>
  <l>Like foolish Prophets forth; their Words to Scorn</l>
  <l>Are scatter'd, and their Mouths are stopt with Dust. </l>
</lg>
```

The <rhyme> element may be used to mark the words (or parts of words) which rhyme according to a predefined pattern:

```
<lg  type="couplet"  rhyme="aa">
  <l>Outside in the distance a wildcat did <rhyme>growl</rhyme>
  </l>
  <l>Two riders were approaching and the wind began to <rhyme>howl</rhyme>
  </l>
</lg>
```

The *label* attribute is used to specify which parts of a rhyme scheme a given set of rhyming words represent:

```
<lg  type="quatrain"  rhyme="abab">
  <l>I wander thro' each charter'd <rhyme  label="a">street</rhyme>,</l>
  <l>Near where the charter'd Thames does <rhyme  label="b">flow</rhyme>,</l>
  <l>And mark in every face I <rhyme  label="a">meet</rhyme>
  </l>
  <l>Marks of weakness, marks of <rhyme  label="b">woe</rhyme>.</l>
</lg>
<lg  rhyme="abab">
  <l>In every cry of every <rhyme  label="a">Man</rhyme>
  </l>
  <l>In every Infant's cry of <rhyme  label="b">fear</rhyme>,</l>
  <l>In every voice, in every <rhyme  label="a">ban</rhyme>,</l>
  <l>The mind-forg'd manacles I <rhyme  label="b">hear</rhyme>.</l>
</lg>
```

Within a given scope, all <rhyme> elements with the same value for their *label* attribute are assumed to rhyme with each other: thus, in the above example, the two rhymes labelled `a` in the first stanza rhyme with each other, but not necessarily with those labelled `a` in the second stanza. The scope is defined by the nearest ancestor element for which the *rhyme* attribute has been supplied.

The <rhyme> element can appear anywhere within a verse line, and not necessarily around a single word. It can thus be used to mark quite complex internal rhyming schemes, as in the following example:

```
<lg  rhyme="ABCCBBA">
  <l>The sunlight on the <rhyme  label="A">garden</rhyme>
  </l>
  <l>
    <rhyme  label="A">Harden</rhyme>s and grows <rhyme  label="B">cold</rhyme>,</l>
  <l>We cannot cage the <rhyme  label="C">minute</rhyme>
  </l>
  <l>Wi<rhyme  label="C">thin it</rhyme>s nets of <rhyme  label="B">gold</rhyme>
  </l>
  <l>When all is <rhyme  label="B">told</rhyme>
  </l>
  <l>We cannot beg for <rhyme  label="A">pardon</rhyme>.</l>
</lg>
```

This mechanism, although reasonably simple for simple cases, may not be appropriate for more complex applications. In general, rhyme may be considered as a special form of 'correspondence', and hence encoded using the mechanisms defined for that purpose in section *17.5. Correspondence and Alignment.* Similar considerations apply to other metrical features such as alliteration or assonance.

To use the correspondence mechanisms to represent the complex rhyming pattern of the above example, each <rhyme> element must be given a unique identifier, as follows:

```
<lg  rhyme="AB-BBA">
  <l>The sunlight on the <rhyme  xml:id="V-A1">garden</rhyme>
  </l>
  <l>
    <rhyme  xml:id="V-A2">Harden</rhyme>s and grows <rhyme  xml:id="V-B1">cold,</rhyme>
  </l>
  <l>We cannot cage the <rhyme  xml:id="V-C1">minute</rhyme>
  </l>
  <l>Wi<rhyme  xml:id="V-C2">thin it</rhyme>s nets of <rhyme  xml:id="V-B2">gold</rhyme>
  </l>
  <l>When all is <rhyme  xml:id="V-B3">told</rhyme>
  </l>
  <l>We cannot beg for <rhyme  xml:id="V-A3">pardon</rhyme>.</l>
</lg>
```

Now that each rhyming word, or part-word, has been tagged and allocated an arbitrary identifier, the general purpose <link> element may be used to indicate which of the <rhyme> elements share the same rhyme, as follows:

```
<linkGrp  type="rhyme">
  <link  target="#V-A1 #V-A2 #V-A3"/>
  <link  target="#V-B1 #V-B2 #V-B3"/>
  <link  target="#V-C1 #V-C2"/>
</linkGrp>
```

For further discussion of the <link> and <linkGrp> element, see section *17.5. Correspondence and Alignment.*

The <rhyme> and <caesura> phrase level elements are made available by the `model.lPart` class when the module defined by this chapter is included in a schema.

## 6.6 Metrical Notation Declaration

When the module defined in this chapter is included in a schema, a specialized element is optionally available in the <encodingDesc> element of the TEI header to document the metrical notation used in marking up a text.

`<metDecl>` (metrical notation declaration) documents the notation employed to represent a metrical pattern when this is specified as the value of a *met*, *real*, or *rhyme* attribute on any structural element of a metrical text (e.g. <lg>, <l>, or <seg>).

`@pattern` (regular expression pattern) specifies a regular expression defining any value that is legal for this notation.

`<metSym>` (metrical notation symbol) documents the intended significance of a particular character or character sequence within a metrical notation, either explicitly or in terms of other <metSym> elements in the same <metDecl>.

`@value` specifies the character or character sequence being documented.

`@terminal` specifies whether the symbol is defined in terms of other symbols (*terminal* is set to false) or in prose (*terminal* is set to true).

As with other components of the header, metrical notation may be specified either formally or informally. In a formal specification, every symbol used in the metrical notation must be documented by a corresponding <metSym> element; in an informal one, only a brief prose description of the way in which the notation is used need be given. In either case, the optional *pattern* attribute may be used to supply a regular expression which a processor can use to validate expressions in the intended notation. The following constraints apply:

- if *pattern* is supplied, any notation used which does not conform to it should be regarded as invalid
- if any <metSym> is defined, then any notation using undefined symbols should be regarded as invalid
- if both pattern and symbol are defined, then every symbol appearing explicitly within pattern must be defined
- symbols which are not matched by *pattern* may be defined within a <metDecl> element

As a simple example, consider the case of the notation in which metrical prominence, metrical feet, and line boundaries are all to be encoded. Legal specifications in this notation may be written for any sequence of metrically prominent or non-prominent features, optionally separated by foot or metrical line boundaries at arbitrary points. Assuming that the symbol *1* is used for metrical prominence, *0* for non-prominence, | for foot boundary and / for line boundary, then the following declaration achieves this objective:

```
<metDecl pattern="((1|0)+\|?/?)*">
  <metSym value="1">metrical prominence</metSym>
  <metSym value="0">metrical non-prominence</metSym>
  <metSym value="|">foot boundary</metSym>
  <metSym value="/">metrical line boundary</metSym>
</metDecl>
```

The same notation might also be specified less formally, as follows:

```
<metDecl>
  <p>Metrically prominent syllables are marked '1' and other
     syllables '0'. Foot divisions are marked by a vertical bar,
     and line divisions with a solidus.</p>
  <p>This notation may be applied to any metrical unit, of any
     size (including, for example, individual feet as well as
     groups of lines).</p>
</metDecl>
```

Note that in this case, because the *pattern* attribute has not been supplied, no processor can validate *met* attribute values within the text which use this metrical notation.

For more complex cases, it will often be more convenient to define a notation incrementally. The *terminal* attribute should be used to indicate for a given symbol whether or not it may be re-defined in terms of other symbols used within the same notation. For example, here is a notation for encoding classical metres, in which symbols are provided for the most common types of foot. These symbols are themselves documented within the same notation, in terms of more primitive long and short syllables:

```
<metDecl pattern="[DTIS3A]+">
  <metSym n="dactyl" value="D"
    terminal="false">-oo</metSym>
  <metSym n="trochee" value="T"
    terminal="false">-o</metSym>
  <metSym n="iamb" value="I" terminal="false">o-</metSym>
  <metSym n="spondee" value="S"
    terminal="false">--</metSym>
  <metSym n="tribrach" value="3"
    terminal="false">ooo</metSym>
  <metSym n="anapaest" value="A"
    terminal="false">oo-</metSym>
  <metSym value="o">short syllable</metSym>
  <metSym value="-">long syllable</metSym>
</metDecl>
```

Note here the use of the global *n* attribute to supply an additional name for the symbols being documented.

Where an encoder wishes to use more than one different pattern for metrical notation, multiple <metDecl> elements may be included in the <encodingDesc>, each supplied with an *xml:id*. The *decls* attribute may be used in the text of the document to specify which <metDecl> is in force at a particular point in the text. In this example, two <metDecl>s are defined in the header, one with an English verse pattern and one with a French pattern. In the body of the document, there are two  elements, one declaring the English pattern and one the French:

```
<encodingDesc>
<!-- ... -->
  <metDecl xml:id="md_en" type="met"
    pattern="((SU|US)USUSUSUS/)">
    <metSym value="S">stressed syllable</metSym>
    <metSym value="U">unstressed syllable</metSym>
    <metSym value="/">metrical line boundary</metSym>
  </metDecl>
  <metDecl xml:id="md_fr" type="met"
    pattern="(AAAAAT\|AAAAT(A)?)">
    <metSym value="T">syllabe tonique</metSym>
    <metSym value="A">syllabe atone</metSym>
    <metSym value="|">pause métrique</metSym>
  </metDecl>
<!-- ... -->
</encodingDesc>
<!-- ... -->
<body>
  <div decls="#md_en">
    <lg>
      <l>
<!-- ... -->
      </l>
<!-- ... -->
    </lg>
  </div>
  <div decls="#md_fr">
    <lg>
      <l>
<!-- ... -->
      </l>
<!-- ... -->
```

```
    </lg>
  </div>
</body>
```

## 6.7 Encoding Procedures for Other Verse Features

A number of procedures that may be of particular concern to encoders of verse texts are dealt with elsewhere in these guidelines. Some aspects of layout and physical appearance, especially important in the case of free verse, are dealt with in chapter *12. Representation of Primary Sources*. Some initial recommendations for the encoding of phonetic or prosodic transcripts, which may be helpful in the analysis of sound structures in poetry, are to be found in chapter *8. Transcriptions of Speech*; it may also be found convenient to use standard entity names (those proposed for the International Phonetic Alphabet suggest themselves) to mark positions of suprasegmentals such as primary and secondary stress, or other aspects of accentual structure.

As already indicated, chapter *17. Linking, Segmentation, and Alignment* contains much which will be found useful for the aligning of multiple levels of commentary and structure within verse analysis. Encoders of verse (as of other types of literary text) will frequently wish to attach identifying labels to portions of text that are not part of a system of hierarchical divisions, may overlap with one another, and/or may be discontinuous; for instance passages associated with particular characters, themes, images, allusions, topoi, styles, or modes of narration. Much of the computerized analysis of verse seems likely to require dividing texts up into blocks in this way. The  element discussed in *18.3. Spans and Interpretations* provides the means for doing this. Finally, the procedures for the tagging of feature structures, described in chapter *19. Feature Structures*, provide a powerful means of encoding a wide variety of aspects of verse literature, including not only the metrical structures discussed above, but also such stylistic and rhetorical features as metaphor.

For other features it must for the time being be left to encoders to devise their own terminology. Elements such as <metaphor tenor="..." vehicle="..."> ... </metaphor> might well suggest themselves; but given the problems of definition involved, and the great richness of modern metaphor theory, it is clear that any such format, if predefined by these Guidelines, would have seemed objectionable to some and excessively restrictive to many. Leaving the choice of tagging terminology to individual encoders carries with it one vital corollary, however: the encoder must be utterly explicit, in the TEI header, about the methods of tagging used and the criteria and definitions on which they rest. Where no formal elements are currently proposed, such information may readily be given as simple prose description within the <encodingDesc> element defined in section *2.3. The Encoding Description*.

## 6.8 Module for Verse

The module described in this chapter makes available the following components:

**Module verse**: **Verse structures**

- *Elements defined*: caesura metDecl metSym rhyme
- *Classes defined*: att.enjamb att.metrical

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 7 Performance Texts

This module is intended for use when encoding printed dramatic texts, screen plays or radio scripts, and written transcriptions of any other form of performance.

Section *7.1. Front and Back Matter* discusses elements such as cast lists, which can appear only in the front or back matter of printed dramatic texts. Section *7.2. The Body of a Performance Text* discusses the structural components of performance texts: these include major structural divisions such as acts and scenes (section *7.2.1. Major Structural Divisions*); individual speeches (section *7.2.2. Speeches and Speakers*); groups of speeches (section *7.2.3. Grouped Speeches*); stage directions (section *7.2.4. Stage Directions*); and the elements making up individual speeches (section *7.2.5. Speech Contents*). Section *7.2.6. Embedded Structures* discusses ways of encoding units which cross the simple hierarchic structure so far defined, such as embedded songs or masques. Finally, section *7.3. Other Types of Performance Text* discusses a small number of additional elements characteristic of screen plays and radio or television scripts, as well as some elements for representing technical stage directions such as lighting or blocking.

The default structure for dramatic texts is similar to that defined by chapter *4. Default Text Structure*, as further discussed in section *7.2.1. Major Structural Divisions*.

Two element classes are used by this module. The `model.frontPart.drama` class supplies specialized elements which can appear only in the front or back matter of performance texts. The `model.stageLike` class supplies a set of elements for stage directions and similar items such as camera movements, which can occur between or within speeches.

## 7.1 Front and Back Matter

In dramatic texts, as in all TEI-conformant documents, the header element is followed by a <text> element, which contains optional front and back matter, and either a <body> or else a <group> of nested <text> elements. For more information on these, see chapter *4. Default Text Structure*.

The <front> and <back> elements are most likely to be of use when encoding preliminary materials in published performance texts. When the module defined by this chapter is included in a schema, the following additional elements not generally found in other forms of text become available as part of the front or back matter:

`<performance>` (performance) contains a section of front or back matter describing how a dramatic piece is to be performed in general or how it was performed on some specific occasion.

`<prologue>` (prologue) contains the prologue to a drama, typically spoken by an actor out of character, possibly in association with a particular performance or venue.

`<epilogue>` (epilogue) contains the epilogue to a drama, typically spoken by an actor out of character, possibly in association with a particular performance or venue.

`<set>` (setting) contains a description of the setting, time, locale, appearance, etc., of the action of a play, typically found in the front matter of a printed performance text (not a stage direction).

`<castList>` (cast list) contains a single cast list or dramatis personae.

Elements for encoding each of these specific kinds of front matter are discussed in the remainder of this section, in the order given above. In addition, the front matter of dramatic texts may include the same elements as that of any other kind of text, notably title pages and various kinds of text division, as discussed in section *4.5. Front Matter*. The encoder may choose to ignore the specialized elements discussed in this section and instead use constructions of the type  or <div1 type="set">.

Most other material in the front matter of a performance text will be marked with the default text structure elements described in chapter *4. Default Text Structure*. For example, the title page, dedication, other commendatory material, preface, etc., in a printed text should be encoded using  or <div1> elements, containing headings, paragraphs, and other core tags.

### 7.1.1 The Set Element

A special form of note describing the setting of a dramatic text (that is, the time and place of its action) is sometimes found in the front matter.

`<set>` (setting) contains a description of the setting, time, locale, appearance, etc., of the action of a play, typically found in the front matter of a printed performance text (not a stage direction).

Descriptions of the setting may also appear as initial stage directions in the body of the play, but such descriptions should be marked as stage directions, not <set>. The <set> element should be used only where the description forms part of the front matter, as in the following examples:

```
<front>
  <castList>
    <castItem> ... </castItem>
  </castList>
  <set>
    <p>The action of the play is set in Chicago's
        Southside, sometime between World War II and the
        present.</p>
  </set>
</front>
```

```
<front>
  <titlePage  type="half-title">
    <docTitle>
      <titlePart>Peer Gynt</titlePart>
    </docTitle>
  </titlePage>
  <div  type="copyright_page"/>
  <div  type="Contents"/>
  <div  type="Introduction"/>
  <div  type="note">
    <head>Note on the Translation</head>
    <p> ... </p>
  </div>
  <div  type="Dramatis_Personae">
    <head>Characters</head>
    <castList>
      <castItem>
<!-- ... -->
      </castItem>
    </castList>
  </div>
  <set>
    <p>The action, which opens in the beginning of the nineteenth
        century, and ends around the 1860s, takes place partly in
        Gudbrandsdalen, and on the mountains around it, partly on the coast
        of Morocco, in the desert of Sahara, in a madhouse at Cairo, at sea,
        etc.</p>
  </set>
  <performance>
    <p>
<!-- ... -->
    </p>
  </performance>
</front>
```

### 7.1.2 Prologues and Epilogues

Many plays in the Western tradition include in their front matter a prologue, spoken by an actor, generally not in character. Similar speeches often also occur at the end of the play, as epilogues. The elements <prologue> and <epilogue> are provided for the encoding of such features within the front or back matter, where appropriate.

`<prologue>` (prologue) contains the prologue to a drama, typically spoken by an actor out of character, possibly in association with a particular performance or venue.

`<epilogue>` (epilogue) contains the epilogue to a drama, typically spoken by an actor out of character, possibly in association with a particular performance or venue.

A prologue may be encoded just like a distinct poem, as in the following example:

```
<front>
  <prologue>
    <head>Prologue, spoken by <name>Mr. Hart</name>
    </head>
    <l>Poets like Cudgel'd Bullys, never do</l>
    <l>At first, or second blow, submit to you;</l>
    <l>But will provoke you still, and ne're have done,</l>
    <l>Till you are weary first, with laying on:</l>
    <l>We patiently you see, give up to you,</l>
    <l>Our Poets, Virgins, nay our Matrons too.</l>
  </prologue>
  <castList>
    <head>The Persons</head>
    <castItem> ... </castItem>
  </castList>
  <set>
    <head>The SCENE</head>
    <p>London</p>
  </set>
</front>
```

A prologue or epilogue may also be encoded as a speech, using the <sp> element described in section *3.13.2. Core Tags for Drama*. This is particularly appropriate where stage directions, etc., are involved, as in the following example:

```
<epilogue>
  <head>Written by <name>Colley Cibber, Esq</name>
     and spoken by <name>Mrs. Cibber</name>
  </head>
  <sp>
    <lg  type="stanza">
      <l>Since Fate has robb'd me of the hapless Youth,</l>
      <l>For whom my heart had hoarded up its truth;</l>
      <l>By all the Laws of Love and Honour, now,</l>
      <l>I'm free again to chuse, — and one of you</l>
    </lg>
    <lg  type="stanza">
      <l>Suppose I search the sober Gallery; — No,</l>
      <l>There's none but Prentices — & Cuckolds all a row:</l>
      <l>And these, I doubt, are those that make 'em so.</l>
    </lg>
    <stage>Pointing to the Boxes.</stage>
    <lg  type="stanza">
      <l>'Tis very well, enjoy the jest:</l>
    </lg>
  </sp>
</epilogue>
```

In cases where the prologue or epilogue is clearly a significant part of the dramatic action, it may be preferable to include it in the body of a text, rather than in the front or back matter. In such cases, the encoder (and theatrical tradition) will determine whether or not to regard it as a new scene or division, or simply the final speech in the play. In the First Folio version of Shakespeare's *Tempest*, for example, Prospero's final speech is clearly marked off as a distinct textual unit by the headings and layout of the page, and might therefore be encoded as back matter:

```
<text>
  <body>
    <div1  type="scene">
      <sp>
        <l  part="Y">I'le deliver all,</l>
        <l>And promise you calme Seas, auspicious gales,</l>
        <l>Be free and fare thou well: please you, draw neere.</l>
        <stage>Exeunt omnes.</stage>
      </sp>
    </div1>
  </body>
  <back>
    <epilogue>
      <head>Epilogue, spoken by Prospero.</head>
      <sp>
        <l>Now my Charmes are all ore-throwne,</l>
        <l>And what strength I have's mine owne</l>
        <l>As you from crimes would pardon'd be,</l>
        <l>Let your Indulgence set me free.</l>
      </sp>
      <stage>Exit</stage>
    </epilogue>
    <set>
      <p>The Scene, an un-inhabited Island.</p>
    </set>
    <castList>
      <head>Names of the Actors.</head>
      <castItem>Alonso, K. of Naples</castItem>
      <castItem>Sebastian, his Brother.</castItem>
      <castItem>Prospero, the right Duke of Millaine.</castItem>
    </castList>
    <trailer>FINIS</trailer>
  </back>
</text>
```

In many modern editions, the editors have chosen to regard Prospero's speech as a part of the preceding scene:

```
<sp>
  <speaker>Prospero</speaker>
  <l  part="Y">I'll deliver all,</l>
  <l>And promise you calm seas, auspicious gales,</l>
  <l>Be free and fare thou well. <stage  type="exit">Exit Ariel</stage>
     Please you, draw near. <stage  type="exit">Exeunt all but Prospero</stage>
    <note  place="margin">Epilogue</note>
  </l>
  <l>Now my charms are all o'erthrown,</l>
  <l>And what strength I have's mine own</l>
  <l>As you from crimes would pardoned be,</l>
  <l>Let your indulgence set me free.</l>
</sp>
<stage  type="mix">He awaits applause, then exit.</stage>
```

### 7.1.3 Records of Performances

Performance texts are not only printed in books to be read, they are also performed. It is common practice therefore to include within the front matter of a printed dramatic text some brief account of particular performances, using the following element:

`<performance>` (performance) contains a section of front or back matter describing how a dramatic piece is to be performed in general or how it was performed on some specific occasion.

The <performance> element may be used to group any and all information relating to the actual performance of a play or screenplay, whether it specifies how the play should be performed in general or how it was performed in practice on some occasion.

Performance information may include complex structures such as cast lists, or paragraphs describing the date and location of a performance, details about the setting portrayed in the performance and so forth. (See the discussion of these specialized structures in section *7.1. Front and Back Matter* above.) If information for more than one performance is being recorded, then more than one <performance> element should be used, wherever possible.

Names of persons, places, and dates of particular significance within the performance record may be explicitly marked using the general purpose <name>, <rs type="place"> and <date> elements described in section *3.6.4. Dates and Times*. No particular elements for such features as stagehouses, directors, etc., are proposed at this time.

For example:

```
<performance>
  <head>Death of a Salesman</head>
  <p>A New Play by Arthur Miller</p>
  <p>Staged by Elia Kazan</p>
  <castList>
    <head>Cast</head>
    <note  rend="small type flush left"
      place="inline">(in order of appearance)</note>
    <castItem>
      <role>Willy Loman</role>
      <actor>Lee J. Cobb</actor>
    </castItem>
    <castItem>
      <role>Linda</role>
      <actor>Mildred Dunnock</actor>
    </castItem>
    <castItem>
      <role>Biff</role>
      <actor>Arthur Kennedy</actor>
    </castItem>
    <castItem>
      <role>Happy</role>
      <actor>Cameron Mitchell</actor>
    </castItem>
<!-- ... -->
  </castList>
  <p>The setting and lighting were designed by
  <name>Jo Mielziner</name>.</p>
  <p>The incidental music was composed by <name>Alex North</name>.</p>
  <p>The costumes were designed by <name>Julia Sze</name>.</p>
  <p>Presented by <name  rend="unmarked">Kermit Bloomgarden</name>
     and <name  rend="unmarked">Walter Fried</name> at the
  <rs  type="place">Morosco Theatre in New York</rs> on
  <date  when="1949-02-10">February 10, 1949</date>.</p>
</performance>
```

Or:

```
<performance>
  <p>La Machine Infernale a été
     représentée pour la première fois au
```

```
  <rs  type="place-theatre">théâtre Louis-Jouvet</rs>
    <rs  type="place-theatre">(Comédie des
        Champs-élysées)</rs>
    <date>le 10 avril 1934</date>,
     avec les décors et les costumes de
  <name>Christian Bérard.</name> ... </p>
</performance>
```

### 7.1.4 Cast Lists

A *cast list* is a specialized form of list, conventionally found at the start or end of a play, usually listing all the speaking and non-speaking roles in the play, often with additional description ('Cataplasma, a maker of Periwigges and Attires') or the name of an actor or actress ('Old Lady Squeamish. Mrs Rutter'). Cast lists may be encoded with the general purpose <list> element described in section *3.8. Lists*, but for more detailed work the following specialized elements are provided:

`<castList>` (cast list) contains a single cast list or dramatis personae.

`<castGroup>` (cast list grouping) groups one or more individual <castItem> elements within a cast list.

`<castItem>` (cast list item) contains a single entry within a cast list, describing either a single role or a list of non-speaking roles.
`@type` characterizes the cast item.

A <castItem> element may contain any mixture of elements taken from the `model.castItemPart` class, members of which (when this module is included) are:

`<role>` (role) contains the name of a dramatic role, as given in a cast list.

`<roleDesc>` (role description) describes a character's role in a drama.

`<actor>` contains the name of an actor appearing within a cast list.

Cast lists often have an internal structure of their own; it is quite usual to find, for example, nobility and commoners, or male and female roles, presented in different groups or sublists. Roles are also often grouped together by their function, for example:

- Sons of Cato:
  - Portius
  - Marcus

A cast list relating to a specific performance may be accompanied by notes about the time or place of that performance, indicating (for example) the name of the theatre where the play was first presented, the name of the producer or director, and so forth. When the cast list relates to a specific performance, it should be embedded within a <performance> element (see section *7.1.3. Records of Performances*), as in the following example:

```
<performance>
  <p>The first performance in Great Britain of <title>Waiting for
        Godot</title> was given at the Arts Theatre, London, on
  <date  when="1955-08-03">3rd August 1955</date>. It was directed by
  <name>Peter Hall</name>, and the décor was by <name>Peter
        Snow</name>. The cast was as follows:</p>
  <castList>
    <castItem>Estragon: Peter Woodthorpe</castItem>
    <castItem>Vladimir: Paul Daneman</castItem>
    <castItem> ... </castItem>
  </castList>
</performance>
```

In this example, the <castItem> elements have no substructure. If desired, however, their components may be more finely distinguished using the elements <role>, <roleDesc>, and <actor>. For example, the second cast item above might be encoded as follows:

```
<castItem>
  <role  xml:id="vlad">Vladimir</role>:
<actor>Paul Daneman</actor>
</castItem>
```

The *ref* attribute on <actor> may be used to associate the name with information about the real-world person identified, as further discussed in section ND. In the previous example, we might associate the name of Paul Daneman with his entry in a widely used bibliography as follows :

```
<actor  ref="https://www.worldcat.org/identities/lccn-n84002994/">Paul Daneman</actor>
```

The global *xml:id* attribute may be used to specify a unique identifier for the <role> element, where it is desired to link speeches within the text explicitly to the role, using the *who* attribute, as further discussed in section *7.2.2. Speeches and Speakers* below.

The occasionally lengthy descriptions of a role sometimes found in written play scripts may be marked using the <roleDesc> element, as in the following example:

```
<castItem>
  <role>Tom Thumb the Great</role>
  <roleDesc>a little hero with a great soul, something violent in his
     temper, which is a little abated by his love for Huncamunca</roleDesc>
  <actor>Young Verhuyk</actor>
</castItem>
```

For non-speaking or un-named roles, a <castItem> may contain a <roleDesc> without an accompanying <role>, for example

```
<castItem>
  <roleDesc>Costermonger</roleDesc>
</castItem>
```

When a list of such minor roles is given together, the *type* attribute of the <castItem> should indicate that it contains more than one role, by taking a value such as list. The encoder may or may not elect to encode each separate constituent within such a composite <castItem>. Thus, either of the following is acceptable:

```
<castItem  type="list">Constables, Drawer, Turnkey, etc.</castItem>
<castItem  type="list">
  <roleDesc>Constables,</roleDesc>
  <roleDesc>Drawer,</roleDesc>
  <roleDesc>Turnkey,</roleDesc>
etc.

</castItem>
```

A group of cast items forming a distinct subdivision of a cast list may be marked as such by using the special purpose <castGroup> element. The *rend* attribute may be used to indicate whether this grouping is indicated in the text by layout alone (i.e. the use of whitespace), by long braces or by some other means. A <castGroup> may contain an optional heading (represented as usual by a <head> element) followed by a series of <castItem> elements:

```
<castGroup  rend="braced">
  <head>friends of Mathias</head>
  <castItem>
    <role>Walter</role>
    <actor>Mr Frank Hall</actor>
  </castItem>
  <castItem>
    <role>Hans</role>
    <actor>Mr F.W. Irish</actor>
  </castItem>
</castGroup>
```

Alternatively, the encoder may prefer to regard the phrase 'friends of Mathias' as a role description, and encode the above example as follows:

```
<castGroup  rend="braced">
  <roleDesc>friends of Mathias</roleDesc>
  <castItem>
    <role>Walter</role>
    <actor>Mr Frank Hall</actor>
  </castItem>
  <castItem>
    <role>Hans</role>
    <actor>Mr F.W. Irish</actor>
  </castItem>
</castGroup>
```

This version has the advantage that all role descriptions are treated alike, rather than in some cases being treated as headings. On the other hand there are also cases, such as the following, where the role description does function more like a heading:

```
<castList>
  <castGroup>
    <head  rend="braced">Mendicants</head>
    <castItem>
      <role>Aafaa</role>
      <actor>Femi Johnson</actor>
    </castItem>
    <castItem>
      <role>Blindman</role>
      <actor>Femi Osofisan</actor>
    </castItem>
    <castItem>
      <role>Goyi</role>
      <actor>Wale Ogunyemi</actor>
    </castItem>
    <castItem>
      <role>Cripple</role>
      <actor>Tunji Oyelana</actor>
    </castItem>
  </castGroup>
  <castItem>
    <role>Si Bero</role>
    <roleDesc>Sister to Dr Bero</roleDesc>
```

```
    <actor>Deolo Adedoyin</actor>
  </castItem>
  <castGroup>
    <head  rend="braced">Two old women</head>
    <castItem>
      <role>Iya Agba</role>
      <actor>Nguba Agolia</actor>
    </castItem>
    <castItem>
      <role>Iya Mate</role>
      <actor>Bopo George</actor>
    </castItem>
  </castGroup>
  <castItem>
    <role>Dr Bero</role>
    <roleDesc>Specialist</roleDesc>
    <actor>Nat Okoro</actor>
  </castItem>
  <castItem>
    <role>Priest</role>
    <actor>Gbenga Sonuga</actor>
  </castItem>
  <castItem>
    <role>The old man</role>
    <roleDesc>Bero's father</roleDesc>
    <actor>Dapo Adelugba</actor>
  </castItem>
</castList>
```

## 7.2 The Body of a Performance Text

The body of a performance text may be divided into structural units, variously called acts, scenes, stasima, entr'actes, etc. All such formal divisions should be encoded using an appropriate text-division element ( , <div1>, <div2>, etc.), as further discussed in section *7.2.1. Major Structural Divisions*. Whether divided up into such units or not, all performance texts consist of sequences of speeches (see *7.2.2. Speeches and Speakers*) and stage directions (see *7.2.4. Stage Directions*). In musical performances, it is also common to identify groups of speeches which act as a single unit, sometimes called a *number*; such units typically float within the structural hierarchy at the same level as speeches preceding or following them and cannot therefore be treated as text-divisions. (see *7.2.3. Grouped Speeches*). Speeches will generally consist of a sequence of *chunk*-level items: paragraphs, verse lines, stanzas, or (in case of uncertainty as to whether something is verse or prose) <ab> elements (see *7.2.5. Speech Contents*).

The boundaries of formal units such as verse lines or paragraphs do not always coincide with speech boundaries. Units such as songs may be discontinuous or shared among several speakers. As described below in section *7.2.6. Embedded Structures*, such fragmentation may be encoded in a relatively simple fashion using the linkage and aggregation mechanisms defined in chapter *17. Linking, Segmentation, and Alignment*.

### 7.2.1 Major Structural Divisions

Large divisions in drama such as acts, scenes, stasima, or entr'actes are indicated by numbered or unnumbered  elements, as described in section *4.1. Divisions of the Body*. The *type* and *n* attributes may be used to define the type of division being marked, and to provide a name or number for it, as in the following example:

```
<body>
  <div1  type="scene"  n="1">
    <head>Night—Faust's Study (i)</head>
  </div1>
```

```
  <div1  type="scene"  n="2">
    <head>Outside the City Gate</head>
  </div1>
</body>
```

Where the largest divisions of a performance text are themselves subdivided, most obviously in the case of plays traditionally divided into acts and scenes, further nested text-division elements may be used, as in this example:

```
<body>
  <div1  type="act"  n="1">
    <head>Act One</head>
    <div2  type="scene"  n="1">
      <stage>Pa Ubu, Ma Ubu</stage>
      <sp>
        <speaker>Pa Ubu</speaker>
        <p>Pschitt!</p>
      </sp>
    </div2>
    <div2  type="scene"  n="2">
      <stage>A room in Pa Ubu's house, where a magnificent
            collation is set out</stage>
    </div2>
  </div1>
  <div1  type="act"  n="2">
    <head>Act Two</head>
    <div2  type="scene"  n="1">
      <head>Scene One</head>
    </div2>
    <div2  type="scene"  n="2">
      <head>Scene Two</head>
    </div2>
  </div1>
</body>
```

In the example above, the <div2> element has been used to represent the 'French scene' convention, (where the entrance of each new set of characters is marked as a distinct unit in the text) and the <div1> element to represent the acts into which the play is divided. The elements chosen are determined only by the hierarchic position of these units in the text as a whole. If the text had no acts, but only scenes, then the scenes might be represented by <div1> elements. Equally, if a play is divided only into 'acts', with no smaller subdivisions, then the <div1> element might be used to represent acts. The *type* should be used, as above, to make explicit the name associated with a particular category of subdivision.

As an alternative to the use of numbered divisions, the encoder may represent all subdivisions with the same element, the unnumbered . The second act in the above example would then be represented as follows:

```
<div  type="act"  n="2">
  <head>Act Two</head>
  <div  type="scene"  n="1">
    <head>Scene One</head>
  </div>
  <div  type="scene"  n="2">
    <head>Scene Two</head>
  </div>
</div>
```

For further discussion of the use of numbered and unnumbered divisions, see section *4.1. Divisions of the Body*.

### 7.2.2 Speeches and Speakers

The following elements are used to identify speeches and speakers in a performance text:

`<sp>` (speech) contains an individual speech in a performance text, or a passage presented as such in a prose or verse text.

`<speaker>` contains a specialized form of heading or label, giving the name of one or more speakers in a dramatic text or fragment.

As noted above, the structure of many performance texts may be analysed as multiply hierarchic: a scene of a verse play, for example, may be divided into speeches and, at the same time, into verse lines. The end of a line may or may not coincide with the end of a speech, and vice versa. Other structures, such as songs, may be discontinuous or split up over several speeches. For some purposes it will be appropriate to regard the verse-structure as the fundamental organizing principle of the text, and for others the speech structure; in some cases, the choice between the two may be arbitrary. The discussion in the remainder of this chapter assumes that it is the speech-based hierarchy which most prominently determines the structure of performance texts, but the same mechanisms could be employed to encode a view of a performance text in which individual speeches were entirely subordinate to the formal units of prose and verse. For more detailed discussion and examples of various treatments of this fundamental issue, refer to chapter *21. Non-hierarchical Structures*.

The *who* attribute and the <speaker> element are both used to indicate the speaker or speakers of a speech, but in rather different ways. The <speaker> element is used to encode the word or phrase actually used within the source text to indicate the speaker: it may contain any string or prefix, and may be thought of as a highly specialized form of stage direction. The *who* attribute however contains one or more pointer values, each of which indicates one or more other XML elements documenting the character to whom the speech is assigned. Typically, this attribute might point to a <person> element in the TEI header *16.2.2. The Participant Description*, to a <role> element in the cast list *7.1.4. Cast Lists*, or even to some external source such as an online handbook of dramatic roles. The most usual case is that the pointer value supplied (prefixed by a sharp sign) corresponds with the value of an *xml:id* attribute, used elsewhere in the document to identify a particular element, as in the following examples:

```
<castList>
  <castItem>
    <role  xml:id="menae">Menaechmus</role>
  </castItem>
  <castItem>
    <role  xml:id="penic">Peniculus</role>
  </castItem>
</castList>
<sp  who="#menae">
  <speaker>Menaechmus</speaker>
  <l>Responde, adulescens, quaeso, quid nomen tibist?</l>
</sp>
<sp  who="#penic">
  <speaker>Peniculus</speaker>
  <l>Etiam derides, quasi nomen non noveris?</l>
</sp>
<sp  who="#menae">
  <speaker>Menaechmus</speaker>
  <l>Non edepol ego te, quot sciam, umquam ante hunc diem</l>
  <l>Vidi neque novi; ...</l>
</sp>
```

If present, a <speaker> element may only appear as the first part of an <sp> element. The distinction between the <speaker> element and the *who* attribute makes it possible to encode uniformly characters whose names are not indicated in a uniform fashion throughout the play, or characters who appear in disguise, as in the following examples:

```
<castList>
  <castItem>
    <role  xml:id="hh">Henry Higgins</role>
  </castItem>
</castList>
<sp  who="#hh">
  <speaker>The Notetaker</speaker>
  <p> ... </p>
</sp>
```

If the speaker attributions are completely regular (and may thus be reconstructed mechanically from the values given for the *who* attribute), or are of no interest for the encoder of the text (as might be the case with editorially supplied attributions in older texts), then the <speaker> element need not be used; the former example above then might look like this:

```
<castList>
  <castItem>
    <role  xml:id="menaechmus">Menaechmus</role>
  </castItem>
  <castItem>
    <role  xml:id="peniculus">Peniculus</role>
  </castItem>
</castList>
<sp  who="#menaechmus">
  <l>Responde, adulescens, quaeso, quid nomen tibist?</l>
</sp>
<sp  who="#peniculus">
  <l>Etiam derides, quasi nomen non noveris?</l>
</sp>
<sp  who="#menaechmus">
  <l>Non edepol ego te, quot sciam, umquam ante hunc diem</l>
  <l>Vidi neque novi; ...</l>
</sp>
```

More than one identifier may be listed as value for the *who* attribute if the speech is spoken by more than one person, as in the following example:

```
<castList>
  <castItem>
    <role  xml:id="nan">Nano</role>
  </castItem>
  <castItem>
    <role  xml:id="cas">Castrone</role>
  </castItem>
</castList>
<stage>Nano and Castrone sing</stage>
<sp  who="#nan #cas">
  <l>Fools, they are the only nation</l>
  <l>Worth men's envy or admiration</l>
</sp>
```

In the event there is a speech that is assigned to a character that is not listed in the source cast list, a <castList> may be encoded inside the <standOff> element to provide an element to which the *who* of <sp> may point.

The <sp> and <speaker> elements are both declared within the core module (see section *3.13. Passages of Verse or Drama*).

### 7.2.3 Grouped Speeches

This module makes available the following additional element for handling groups of speeches:

`<spGrp>` (speech group) contains a group of speeches or songs in a performance text presented in a source as constituting a single unit or 'number'.

The <spGrp> element is intended for cases where the characters in a performance launch into something which might be regarded almost as a kind of separate structural division, typically associated with its own heading or numbering system, but which 'floats' in the text, at the same hierarchic level as speeches preceding or following it. Such units are often numbered, titled, and visually presented as distinct objects within the text. Here is a typical example from a well-known American musical comedy:

```
<spGrp  type="number"  n="3">
  <head>By Strauss : performed by Georges Guetary, Gene Kelly, and Oscar
     Levant</head>
  <sp>
    <speaker>HENRI BAUREL</speaker>
    <lg>
      <l>The waltzes of Mittel Europa </l>
      <l>They charm you and warm you within </l>
      <l>While each day discloses </l>
      <l>What Broadway composes </l>
      <l>Is emptiness pounding on tin.</l>
    </lg>
  </sp>
  <sp  xml:lang="de">
    <speaker>JERRY MULLIGAN: ADAM COOK:</speaker>
    <lg>
      <l>Mein Herr! </l>
      <l>Mein Herr!</l>
      <l>Bitte, bitte!</l>
      <l>Denke, danke!</l>
      <l>Aufwiedersehen! Aufwiedersehen!</l>
    </lg>
  </sp>
  <sp>
    <speaker>HENRI BAUREL:</speaker>
    <lg>
      <l>How can I be civil </l>
      <l>When hearing this drivel? </l>
      <l>It's only for night-clubbing souses. </l>
      <l>Oh give me the free 'n easy </l>
      <l>Waltz that is Viennes-y </l>
      <l>And go tell the band</l>
      <l>If they want a hand</l>
      <l>The waltz must be Strauss's</l>
    </lg>
  </sp>
  <sp>
    <speaker>ALL</speaker>
    <lg>
      <l>Ya ya ya </l>
      <l>Give me oom pah pah...</l>
    </lg>
  </sp>
<!-- ... -->
</spGrp>
```

### 7.2.4 Stage Directions

Both between and within the speeches of a written performance text, it is normal practice to include a wide variety of descriptive directions to indicate non-verbal action. The following elements are provided to represent these:

`<stage>` (stage direction) contains any kind of stage direction within a dramatic text or fragment.

- `@type` indicates the kind of stage direction.

`<move>` (movement) marks the actual movement of one or more characters.

- `@type` characterizes the movement, for example as an entrance or exit.
- `@where` specifies the direction of a stage movement.
- `@perf` (performance) identifies the performance or performances in which this movement occurred as specified by pointing to one or more <performance> elements.

A satisfactory typology of stage directions is difficult to define. Certain basic types such as 'entrance', 'exit', 'setting', 'delivery', are easily identified. But the list is not a closed one, and it is not uncommon to mix types within a single direction. No closed set of values for the *type* attribute is therefore proposed at the present time, though some suggested values are indicated in the list below, which also indicates the range of possibilities.

```
<stage type="setting">The throne descends.</stage>
<stage type="setting">Music</stage>
<stage type="entrance">Enter Husband as being thrown off his horse.</stage>
<stage type="exit">Exit pursued by a bear.</stage>
<stage type="business">He quickly takes the stone out.</stage>
<stage type="delivery">To Lussurioso.</stage>
<stage type="delivery">Aside.</stage>
<stage type="delivery">Not knowing what to say.</stage>
<stage type="costume">Disguised as Ansaldo.</stage>
<stage type="location">At a window.</stage>
<stage type="novelistic">Having had enough, and embarrassed
for the family.</stage>
```

The meaning of the values used for the *type* attribute on <stage> elements may be defined within the <tagUsage> element of the TEI header (described in section *2.3.4. The Tagging Declaration*). For example:

```
<tagUsage gi="stage">This element is used for all stage directions,
editorial or authorial. The type attribute on this element takes
one or more of the following values:
<list type="gloss">
   <label>setting</label>
   <item>describes the set</item>
   <label>blocking</label>
   <item>describes movement across stage, position, etc.</item>
   <label>business</label>
   <item>describes movement other than blocking</item>
   <label>delivery</label>
   <item>describes how the line is said</item>
   <label>motivation</label>
   <item>describes character's emotional state or through line</item>
  </list>
</tagUsage>
```

This approach is purely documentary; in a real project it would generally be more effective to define the range of permitted values explicitly within the project's schema specification, using the techniques described in chapter *24.3. Customization*. For example, a specification like the following might be used to produce a schema in which the *type* attribute of the <stage> element is permitted to take only the values listed above:

```
<schemaSpec ident="myDrama">
  <moduleRef key="core"/>
  <moduleRef key="tei"/>
  <moduleRef key="structure"/>
  <moduleRef key="header"/>
  <moduleRef key="drama"/>
  <elementSpec ident="stage" mode="change">
    <attList>
      <attDef ident="type" mode="replace">
        <valList type="closed">
          <valItem ident="setting">
            <desc>describes the set</desc>
          </valItem>
          <valItem ident="blocking">
            <desc>describes movement across stage, position, etc.</desc>
          </valItem>
          <valItem ident="business">
            <desc>describes movement other than blocking</desc>
          </valItem>
          <valItem ident="delivery">
            <desc>describes how the line is said</desc>
          </valItem>
          <valItem ident="motivation">
            <desc>describes character's emotional state or through line</desc>
          </valItem>
        </valList>
      </attDef>
    </attList>
  </elementSpec>
</schemaSpec>
```

The <stage> element may appear both between and within <sp> elements. It may contain a mixture of phrase level elements, possibly combined into paragraphs, as in the following example:

```
<div1 n="1" type="act">
  <stage type="setting">
    <p>Scene. — A room furnished comfortably and
        tastefully but not extravagantly ...
        The floor is carpeted and a fire burns in the stove.
        It is winter.</p>
    <p>A bell rings in the hall; shortly afterwards the
        door is heard to open. Enter NORA humming a tune ...</p>
  </stage>
  <sp>
    <speaker>Nora</speaker>
    <p>Hide the Christmas Tree carefully, Helen. Be sure the
        children do not see it till this evening, when it is
        dressed. <stage type="delivery">To the PORTER taking
           out her purse</stage> How much?</p>
  </sp>
</div1>
```

The <stage> element may also be used in non-theatrical texts, to mark sound effects or musical effects, etc., as further discussed in section *7.3. Other Types of Performance Text*.

The <move> element is intended to help overcome the fact that the stage directions of a printed text may often not provide full information about either the intended or the actual movement of actors on stage. It may be used to keep

track of entrances and exits in detail, so as to know which characters are on stage at which time. Its attributes permit a relatively formal specification for movements of characters, using user-defined codes to identify the characters involved (the *who* attribute), the direction of the movement (*type* attribute), and optionally which part of the stage is involved (*where* attribute). For stage-historical purposes, a *perf* attribute is also provided; this allows the recording of different <move> elements as taken in different performances of the same text.

The <move> element should be located at the position in the text where the move is presumed to take place. This will often coincide with a stage direction, as in the following simple example:

```
<castList>
  <castItem>
    <role xml:id="bella">Bellafront</role>
  </castItem>
</castList>
<stage type="entrance">
  <move who="#bella" type="enter"/>
Enter Bellafront mad.
</stage>
```

The <move> element can however appear independently of a stage direction, as in the following example:

```
<castList>
  <castItem>
    <role xml:id="lm">Lady Macbeth</role>
  </castItem>
  <castItem>
    <role xml:id="g1">First Gentleman</role>
  </castItem>
<!-- ... -->
</castList>
<sp who="#g1">
  <speaker>Gent.</speaker>
  <p>Neither to you, nor any one; having no witness
     to confirm my speech. <move who="#lm" type="enter" where="C"/>
     Lo you! here she comes. This is her very guise; and,
     upon my life, fast asleep.</p>
</sp>
```

### 7.2.5 Speech Contents

The actual speeches of a dramatic text may be composed of running text, which must be formally organized into paragraphs, in the case of prose (see section *3.1. Paragraphs*), verse lines or line groups in that of verse (see section *3.13. Passages of Verse or Drama*), or <seg> elements, in case of doubt as to whether the material should be treated as verse or prose. The following elements, all of which are defined in the core, are particularly useful when marking units of prose or verse within speeches:

`<p>` (paragraph) marks paragraphs in prose.

`<lb>` (line beginning) marks the beginning of a topographic line in some edition or version of a text.

`<l>` (verse line) contains a single, possibly incomplete, line of verse.

`<lg>` (line group) contains one or more verse lines functioning as a formal unit, e.g. a stanza, refrain, verse paragraph, etc.

Like other milestone elements, the element <lb> additionally bears the attributes *ed* and *edRef*, from its membership in the class `att.edition`:

`att.edition` provides attributes identifying the source edition from which some encoded feature derives.

- `@ed` (edition) supplies a sigil or other arbitrary identifier for the source edition in which the associated feature (for example, a page, column, or line beginning) occurs at this point in the text.

`@edRef` (edition reference) provides a pointer to the source edition in which the associated feature (for example, a page, column, or line beginning) occurs at this point in the text.

As a member of the classes `att.typed` and `att.divLike`, the <lg> element also bears the following attributes:

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.

`@subtype` (subtype) provides a sub-categorization of the element, if needed.

`att.divLike` provides attributes common to all elements which behave in the same way as divisions.

`@org` (organization) specifies how the content of the division is organized.

`@sample` indicates whether this division is a sample of the original source and if so, from which part.

`att.fragmentable` provides attributes for representing fragmentation of a structural element, typically as a consequence of some overlapping hierarchy.

`@part` specifies whether or not its parent element is fragmented in some way, typically by some other overlapping structure: for example a speech which is divided between two or more verse stanzas, a paragraph which is split across a page division, a verse line which is divided between two speakers.

When the `verse` module is included in a schema, the elements <l> and <lg> also gain additional attributes through their membership of the class `att.metrical`:

`att.metrical` defines a set of attributes that certain elements may use to represent metrical information.

`@met` (metrical structure, conventional) contains a user-specified encoding for the conventional metrical structure of the element.

`@rhyme` (rhyme scheme) specifies the rhyme scheme applicable to a group of verse lines.

In many texts, prose and verse may be inextricably mingled; particularly in earlier printed texts, prose may be printed as verse or verse as prose, or it may be impossible to distinguish the two. In cases of doubt, an encoder may prefer to tag the dubious material consistently as verse, to tag it all as prose, to follow the typography of the source text, or to use the neutral <ab> element to contain the speech itself. When this question arises, the <tagUsage> element in the <encodingDesc> element of the header may be used to record explicitly what policy has been adopted.

Even where they can reliably be distinguished, a single speech may frequently contain a mixture of prose (marked as <p>) and verse (marked as <l> or—if stanzaic—<lg>).

The *part* attribute which both <l> and <lg> elements inherit from the `att.fragmentable` class provides one simple way of indicating where the boundaries of a speech and of a verse line or line group do not coincide. The encoder may simply indicate that a line or line group is metrically incomplete by specifying the value Y or N, as in the following example:

```
<sp>
  <speaker>Face</speaker>
  <l  part="N">You most
     notorious whelp, you insolent slave</l>
  <l  part="Y">Dare you do this?</l>
</sp>
<sp>
  <speaker>Subtle</speaker>
  <l  part="Y">Yes faith, yes faith.</l>
</sp>
<sp>
  <speaker>Face</speaker>
  <l  part="Y">Why! Who</l>
  <l  part="Y">Am I, my mongrel? Who am I?</l>
</sp>
```

```
<sp>
  <speaker>Subtle</speaker>
  <l  part="Y">I'll tell you,</l>
<!-- ... -->
</sp>
```

Alternatively, where the fragments of the line or line group are consecutive in the text (though possibly interrupted by stage directions), the values I (initial), M (medial), and F (final) may be used to indicate how metrical lines are constituted:

```
<sp>
  <speaker>Face</speaker>
  <l>You most
     notorious whelp, you insolent slave</l>
  <l  part="I">Dare you do this?</l>
</sp>
<sp>
  <speaker>Subtle</speaker>
  <l  part="M">Yes faith, yes faith.</l>
</sp>
<sp>
  <speaker>Face</speaker>
  <l  part="F">Why! Who</l>
  <l  part="I">Am I, my mongrel? Who am I?</l>
</sp>
<sp>
  <speaker>Subtle</speaker>
  <l  part="F">I'll tell you,</l>
<!-- ... -->
</sp>
```

In dramatic texts, the <lg> or line group element is most often of use for the encoding of songs and other stanzaic material. Line groups may be fragmented across speakers in the same way as individual lines, and the same set of attributes may be used to record this fact. The element <spGrp> is provided in order to simplify the situation, very common in performances, where performance of a single entity, such as a song, is shared amongst several performers, as in the following example:

```
<spGrp>
  <head>Song — Sir Joseph</head>
  <sp>
    <l>I am the monarch of the sea,</l>
    <l>The ruler of the Queen's Navee.</l>
    <l>Whose praise Great Britain loudly chants.</l>
  </sp>
  <sp>
    <speaker>Cousin Hebe</speaker>
    <l>And we are his sisters and his cousins and his aunts!</l>
  </sp>
  <sp>
    <speaker>Rel.</speaker>
    <l>And we are his sisters and his cousins and his aunts!</l>
  </sp>
<!-- ... -->
</spGrp>
```

This encoding however does not indicate that the three lines of Sir Joseph's song and the two lines following it together constitute a single verse stanza. This can be indicated by using the *part* attribute, as follows:

```
<spGrp>
  <head>Song — Sir Joseph</head>
  <sp>
    <lg  part="I">
      <l>I am the monarch of the sea,</l>
      <l>The ruler of the Queen's Navee.</l>
      <l>Whose praise Great Britain loudly chants.</l>
    </lg>
  </sp>
  <sp>
    <speaker>Cousin Hebe</speaker>
    <lg  part="M">
      <l>And we are his sisters and his cousins and his aunts!</l>
    </lg>
  </sp>
  <sp>
    <speaker>Rel.</speaker>
    <lg  part="F">
      <l>And we are his sisters and his cousins and his aunts!</l>
    </lg>
  </sp>
<!-- ... -->
</spGrp>
```

### 7.2.6 Embedded Structures

Although primarily composed of speeches, performance texts often contain other structural units such as songs or strophes which are shared among different speakers. More generally, complex nested structures of plays within plays, interpolated masques, or interludes are far from uncommon. In more modern material, comparably complex structural devices such as flashback or nested playback are equally frequent. In all kinds of performance material, it may be necessary to indicate several actions which are happening simultaneously.

A number of different devices are available within the TEI scheme to support these complexities in the general case. Texts may be composite or self-nesting (see section *4.3.1. Grouped Texts*) and multiple hierarchies may be defined (see chapter *21. Non-hierarchical Structures*). The TEI encoding scheme provides a variety of linking mechanisms, which may be used to indicate temporal alignment and aggregation of fragmented structures. In this section we provide a few specific examples of the application of these techniques to performance texts:

- the use of the <floatingText> element
- the use of the *part* attribute on fragmentary <lg> elements
- the use of the *next* and *prev* attributes on fragments of embedded structures to join them into a larger whole
- the use of the <join> element to define a 'virtual element' composed of the fragments indicated

When the whole of a song appears within a single speech, it may require no special treatment if it is considered to form a part of the speech:

```
<sp>
  <speaker>Kelly</speaker>
  <stage>(calmly).</stage>
  <p>Aha, so you've bad minds along with th' love of gain.
     You thry to pin on others th' dirty decorations that
     may be hangin' on your own coats.</p>
  <stage>(He points, one after the other at Conroy, Bull,
     and Flagonson. Lilting)</stage>
```

```
  <lg  type="song">
    <l>Who were you with last night?</l>
    <l>Who were you with last night?</l>
    <l>Will you tell your missus when you go home</l>
    <l>Who you were with last night?</l>
  </lg>
</sp>
<sp>
  <speaker>Flagonson</speaker>
  <stage>(in anguished indignation).</stage>
  <p>This is more than a hurt to us: this hits at the
     decency of the whole nation!</p>
</sp>
```

If however, the song is to be regarded as forming a distinct item, perhaps with its own front and back matter, it may be better to regard it as a floating text:

```
<sp>
  <speaker>Kelly</speaker>
  <stage>(calmly).</stage>
  <p>Aha, so you've bad minds along with ...</p>
</sp>
<stage>(He points, one after the other at Conroy, Bull,
and Flagonson. Lilting):</stage>
<floatingText>
  <front>
    <titlePart>Kelly's Song</titlePart>
  </front>
  <body>
    <l>Who were you with last night?</l>
    <l>Who were you with last night?</l>
    <l>Will you tell your missus when you go home</l>
    <l>Who you were with last night?</l>
  </body>
</floatingText>
```

When an embedded structure extends across more than one <sp> element, each of its constituent parts must be regarded as a distinct fragment; the problem then facing the encoder is to reconstitute the interrupted whole in some way.

As already noted above, the <spGrp> element may be used to group together consecutive speeches which are grouped together in some way, for example constituting a single song. Alternatively the *part* attribute, typically used to indicate that an <l> element contains a partial, not a complete, verse line, may also be used on the <lg> element, to indicate that the line group is partial rather than complete, thus:

```
<sp>
  <speaker>Kelly</speaker>
  <stage>(wheeling quietly in his semi-dance,
     as he goes out):</stage>
  <lg  type="stanza"  part="I">
    <l>Goodbye to holy souls left here,</l>
    <l>Goodbye to man an' fairy;</l>
  </lg>
</sp>
<sp>
```

```
  <speaker>Widda Machree</speaker>
  <stage>(wheeling quietly in her semi-dance,
     as she goes out):</stage>
  <lg  type="stanza"  part="F">
    <l>Goodbye to all of Leicester Square,</l>
    <l>An' the long way to Tipperary.</l>
  </lg>
</sp>
```

When the fragments of a song are separated by other intervening dialogue, or even when not, they may be linked together with the *next* and *prev* attributes defined in section *17.7. Aggregation.* For example, the line groups making up Ophelia's song might be encoded as follows:

```
<div1  n="4"  type="act">
  <div2  n="5"  type="scene">
    <stage>Elsinore. A room in the Castle.</stage>
    <stage  type="setting">Enter Ophelia, distracted.</stage>
    <sp>
      <speaker>Ophelia</speaker>
      <p>Where is the beauteous Majesty of Denmark?</p>
    </sp>
    <sp>
      <speaker>Queen</speaker>
      <p>How now, Ophelia?</p>
    </sp>
    <sp>
      <speaker>Ophelia</speaker>
      <stage>Singing</stage>
      <lg  next="#Tl2"  xml:id="Tl1"  type="song"
        part="Y">
        <l>How should I your true-love know</l>
        <l>From another one?</l>
        <l>By his cockle hat and staff</l>
        <l>And his sandal shoon.</l>
      </lg>
    </sp>
    <sp>
      <speaker>Queen</speaker>
      <p>Alas, sweet lady, what imports this song?</p>
    </sp>
    <sp>
      <speaker>Ophelia</speaker>
      <p>Say you? Nay, pray you mark.</p>
      <stage>Sings</stage>
      <lg  prev="#Tl1"  xml:id="Tl2"  type="song"
        part="Y">
        <l>He is dead and gone, lady,</l>
        <l>He is dead and gone;</l>
        <l>At his head a grass-green turf,</l>
        <l>At his heels a stone.</l>
      </lg>
      <p>O, ho!</p>
    </sp>
  </div2>
</div1>
```

The *next* and *prev* attributes are discussed in section *17.7. Aggregation*: they form part of the module for alignment and linking; this module must therefore be included in a schema if they are to be used, as further discussed in section *1.2. Defining a TEI Schema*.

The fragments of Ophelia's song might also be linked together using the <join> mechanism described in section *17.7. Aggregation*. The <join> element is specifically intended to encode the fact that several discontiguous elements of the text together form one 'virtual' element. Using this mechanism, the example might be encoded as follows:

```
<text>
  <body>
    <div1 n="4" type="act">
      <div2 n="5" type="scene">
        <stage type="setting">Elsinore. A room in the Castle.</stage>
        <sp>
          <speaker>Queen</speaker>
          <p>How now, Ophelia?</p>
        </sp>
        <sp>
          <speaker>Ophelia</speaker>
          <stage type="delivery">Singing</stage>
          <lg xml:id="TL1" type="song" part="Y">
            <l>How should I your true-love know</l>
            <l>From another one?</l>
            <l>By his cockle hat and staff</l>
            <l>And his sandal shoon.</l>
          </lg>
        </sp>
        <sp>
          <speaker>Queen</speaker>
          <p>Alas, sweet lady, what imports this song?</p>
        </sp>
        <sp>
          <speaker>Ophelia</speaker>
          <p>Say you? Nay, pray you mark.</p>
          <stage type="delivery">Sings</stage>
          <lg xml:id="TL2" type="song" part="Y">
            <l>He is dead and gone, lady,</l>
            <l>He is dead and gone;</l>
            <l>At his head a grass-green turf,</l>
            <l>At his heels a stone.</l>
          </lg>
          <p>O, ho!</p>
          <join type="lg" target="#TL1 #TL2"/>
        </sp>
      </div2>
    </div1>
  </body>
</text>
```

The location of the <join> element is not significant; here it has been placed shortly after the conclusion of the song, in order to have it close to the fragments it unifies.

Like the *next* and *prev* attributes, the <join> element requires the additional module for linking, which is selected as shown above.

### 7.2.7 Simultaneous Action

In printed or written versions of performance texts, a variety of techniques may be used to indicate the temporal alignment of speeches or actions. Speeches may be printed vertically aligned on the page, or braced together; stage

directions (e.g. 'Speaking at the same time') are also often used. In operatic or musical works in particular, the need to indicate timing and alignment of individual parts of a song may lead to very complex layout.

One simple method of indicating the temporal alignment of speeches or actions is to use the <spGrp> element discussed in section *7.2.3. Grouped Speeches*, with a *type* attribute to specify the reason for grouping, as in the following example:

```
<sp>
  <speaker>Mangan</speaker>
  <stage  type="delivery">wildly</stage>
  <p>Look here: I'm going to take off all my clothes.</p>
  <stage  type="action">he begins tearing off his coat.</stage>
</sp>
<spGrp  type="simultaneous"  rend="braced">
  <sp>
    <speaker>Lady Utterword</speaker>
    <p>Mr Mangan!</p>
  </sp>
  <sp>
    <speaker>Captain Shotover</speaker>
    <p>Whats that?</p>
  </sp>
  <sp>
    <speaker>Hector</speaker>
    <p>Ha! ha! Do. Do.</p>
  </sp>
  <sp>
    <speaker>Ellie</speaker>
    <p>Please dont.</p>
  </sp>
  <stage  type="delivery">in consternation</stage>
</spGrp>
<sp>
  <speaker>Mrs. Hushabye</speaker>
  <stage  type="action">catching his arm and stopping him</stage>
  <p>Alfred: for shame! Are you mad?</p>
</sp>
```

In the original, the stage direction 'in consternation' is printed opposite a brace grouping all four speeches, indicating that all four characters speak at once, and that the stage direction applies to all of them. Rather than attempting to represent the appearance of the source, this example encoding represents its presumed meaning: the <stage> element is placed arbitrarily after the last relevant speech, and the four speeches with which it is to be associated are grouped by means of the <spGrp> element. The *rend* attribute is used to specify the fact that the three speeches were grouped by the brace in the copy text. Producing a readable version of the text which simulates the original printed effect may however require more complex markup and processing.

More powerful and more precise mechanisms for temporal alignment are defined in chapter *8. Transcriptions of Speech*. These would be appropriate for encodings the focus of which is on the actual performance of a text rather than its structure or formal properties. The module described in that chapter includes a large number of other detailed proposals for the encoding of such features as voice quality, prosody, etc., which might be relevant to such a treatment of performance texts.

## 7.3 Other Types of Performance Text

Most of the elements and structures identified thus far are derived from traditional theatrical texts. Although other performance texts, such as screenplays or radio scripts, have not been discussed specifically, they can be encoded using the elements and structures listed above. Encoders may however find it convenient to use, as well, the additional specialized elements discussed in this section. For scripts containing very detailed technical information, the <tech> element discussed in section *7.3.1. Technical Information* may also be useful.

Like other texts, screenplays and television or radio scripts may be divided into text divisions marked with  or <div1>, etc. Within units corresponding with the traditional 'act' and 'scene', further subdivisions or sequences may be identified, composed of individual 'shots', each associated with a single camera angle and setting. Shots and sequences should be encoded using an appropriate text-division element (i.e., a <div3> element if numbered division elements are in use and the next largest unit is a <div2>, or a  element if un-numbered divisions are in use) specifying sequence or shot as the value of the *type* attribute, as appropriate.

It is normal practice in screenplays and radio scripts to distinguish directions concerning camera angles, sound effects, etc., from other forms of stage direction. Such texts also generally include far more detailed specifications of what the audience actually sees: descriptions of actions and background, etc. Scripts derived from cinema and television productions may also include texts displayed as captions superimposed on the action. All of these may be encoded using the general purpose <stage> element discussed in section *7.2.4. Stage Directions*, and distinguished by means of its *type* attribute. Alternatively, or in addition, the following more specific elements may be used, where clear distinctions can be made:

`<view>` (view) describes the visual context of some part of a screen play in terms of what the spectator sees, generally independent of any dialogue.

`<camera>` (camera) describes a particular camera angle or viewpoint in a screen play.

`<caption>` (caption) contains the text of a caption or other text displayed as part of a film script or screenplay.

`<sound>` (sound) describes a sound effect or musical sequence specified within a screen play or radio script.

`@type` categorizes the sound in some respect, e.g. as music, special effect, etc.

`@discrete` indicates whether the sound overlaps the surrounding speeches or interrupts them.

Some examples of the use of these elements follow:

```
<camera>Angle on Olivia.</camera>
<view>Ryan's wife, standing nervously alone on the sidelines,
biting her lip. She's scared and she shows it.</view>
```

Where particular words or phrases within a direction are emphasized (by change of typeface or use of capital letters), an appropriate phrase-level element may be used to indicate the fact, as in the following examples, where certain words in the original are given in small capitals:

```
<view>George glances at the window—and freezes.
<camera>New angle—shock cut</camera> Out the window
the body of a dead man suddenly slams into
<hi>frame</hi>. He dangles grotesquely,
held up by his coat caught on a protruding bolt.
George gasps. The train <hi>whistle</hi> screams.</view>
```

```
<view>Ext. TV control van—Early morning.
The <name>T.V. announcer</name> from the Ryan interview
stands near the Control Van, the lake in b.g.</view>
<sp>
  <speaker>T.V. Announcer</speaker>
  <p>Several years ago, Jack Ryan was a highly
     successful hydroplane racer ...</p>
</sp>
```

All of these elements, like other stage directions, can appear both within and between speeches.

```
<sp>
  <speaker>TV Announcer VO</speaker>
  <p>Working with Ryan are his two coworkers—
     Strut Bowman, the mechanical engineer—
  <view>
      <camera>Angle on Strut</camera>
         standing in the tow boat, walkie-talkie in hand,
         watching Ryan carefully.</view>
     —and Roger Dalton, a rocket
     systems analyst, and one of the scientists
     from the Jet Propulsion Lab ...</p>
</sp>
```

```
<sp>
  <speaker>Benjy</speaker>
  <p>Now to business.</p>
</sp>
<sp>
  <speaker>Ford and Zaphod</speaker>
  <p>To business.</p>
</sp>
<sound>Glasses clink.</sound>
<sp>
  <speaker>Benjy</speaker>
  <p>I beg your pardon?</p>
</sp>
<sp>
  <speaker>Ford</speaker>
  <p>I'm sorry, I thought you were proposing a toast.</p>
</sp>
```

```
<camera>Zoom in to overlay showing some stock film
of hansom cabs galloping past.</camera>
<caption>London, 1895.</caption>
<caption>The residence of Mr Oscar Wilde.</caption>
<sound>Suitably classy music starts.</sound>
<view>Mix through to Wilde's drawing room. A crowd of suitably
dressed folk are engaged in typically brilliant conversation,
laughing affectedly and drinking champagne.</view>
<sp>
  <speaker>Prince of Wales</speaker>
  <p>My congratulations, Wilde. Your latest play is a great success.</p>
</sp>
```

### 7.3.1 Technical Information

Traditional stage scripts may contain additional technical information about such production-related factors as lighting, 'blocking' (that is, detailed notes on actors' movements), or props required at particular points. More technical information about intended production effects may also appear in published versions of screenplays or movie scripts. Where these are presented simply as marginal notes, they may be encoded using the general-purpose <note> element defined in section *3.9. Notes, Annotation, and Indexing*. Alternatively, they may be formally distinguished from other stage directions by using the specialized <tech> element:

`<tech>` (technical stage direction) describes a special-purpose stage direction that is not meant for the actors.

`@type` categorizes the technical stage direction.

`@perf` (performance) points to one or more <performance> elements documenting the performance or performances to which this technical direction applies.

Like stage directions, <tech> elements can appear anywhere within a speech or between speeches.

## 7.4 Module for Performance Texts

The module described in this chapter makes available the following components:

**Module drama**: **Performance texts**

- *Elements defined*: actor camera caption castGroup castItem castList epilogue move performance prologue role roleDesc set sound spGrp tech view

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 8 Transcriptions of Speech

The module described in this chapter is intended for use with a wide variety of transcribed spoken material. It should be stressed, however, that the present proposals are not intended to support unmodified every variety of research undertaken upon spoken material now or in the future; some discourse analysts, some phonologists, and doubtless others may wish to extend the scheme presented here to express more precisely the set of distinctions they wish to draw in their transcriptions. Speech regarded as a purely acoustic phenomenon may well require different methods from those outlined here, as may speech regarded solely as a process of social interaction.

This chapter begins with a discussion of some of the problems commonly encountered in transcribing spoken language (section *8.1. General Considerations and Overview*). Section *8.2. Documenting the Source of Transcribed Speech* documents some additional TEI header elements which may be used to document the recording or other source from which transcribed text is taken. Section *8.3. Elements Unique to Spoken Texts* describes the basic structural elements provided by this module. Finally, section *8.4. Elements Defined Elsewhere* of this chapter reviews further problems specific to the encoding of spoken language, demonstrating how mechanisms and elements discussed elsewhere in these Guidelines may be applied to them.

## 8.1 General Considerations and Overview

There is great variation in the ways different researchers have chosen to represent speech using the written medium.[67] This reflects the special difficulties which apply to the encoding or *transcription* of speech. Speech varies according to a large number of dimensions, many of which have no counterpart in writing (for example, tempo, loudness, pitch, etc.). The audibility of speech recorded in natural communication situations is often less than perfect, affecting the accuracy of the transcription. Spoken material may be transcribed in the course of linguistic, acoustic, anthropological, psychological, ethnographic, journalistic, or many other types of research. Even in the same field, the interests and theoretical perspectives of different transcribers may lead them to prefer different levels of detail in the transcript and different styles of visual display. The production and comprehension of speech are intimately bound up with the situation in which speech occurs, far more so than is the case for written texts. A speech transcript must therefore include some contextual features; determining which are relevant is not always simple. Moreover, the ethical problems in recording and making public what was produced in a private setting and intended for a limited audience are more frequently encountered in dealing with spoken texts than with written ones.

Speech also poses difficult structural problems. Unlike a written text, a speech event takes place in time. Its beginning and end may be hard to determine and its internal composition difficult to define. Most researchers agree that the utterances or *turns* of individual speakers form an important structural component in most kinds of speech, but these are rarely as well-behaved (in the structural sense) as paragraphs or other analogous units in written texts: speakers frequently interrupt each other, use gestures as well as words, leave remarks unfinished and so on. Speech itself, though it may be represented as words, frequently contains items such as vocalized pauses which, although only semi-lexical, have immense importance in the analysis of spoken text. Even non-vocal elements such as gestures may be regarded as forming a component of spoken text for some analytic purposes. Below the level of the individual utterance, speech may be segmented into units defined by phonological, prosodic, or syntactic phenomena; no clear agreement exists, however, even as to appropriate names for such segments.

Spoken texts transcribed according to the guidelines presented here are organized as follows. The overall structure of a TEI spoken text is identical to that of any other TEI text: the <TEI> element for a spoken text contains a <teiHeader> element, followed by a <text> element. Even texts primarily composed of transcribed speech may also include conventional front and back matter, and may even be organized into divisions like printed texts.

We may say, therefore, that these Guidelines regard transcribed speech as being composed of arbitrary high-level units called *texts*. A spoken <text> might typically be a conversation between a small number of people, a lecture, a broadcast TV item, or a similar event. Each such unit has associated with it a <teiHeader> providing detailed contextual information such as the source of the transcript, the identity of the participants, whether the speech is scripted or spontaneous, the physical and social setting in which the discourse takes place and a range of other aspects. Details of the header in general are provided in chapter *2. The TEI Header*; the particular elements it provides for use with spoken texts are described below (*8.2. Documenting the Source of Transcribed Speech*). Details concerning additional elements which may be used for the documentation of participant and contextual information are given in *16.2. Contextual Information*.

[67] For a discussion of several of these see ?; ?; and ?.

Defining the bounds of a spoken text is frequently a matter of arbitrary convention or convenience. In public or semi-public contexts, a text may be regarded as synonymous with, for example, a *lecture*, a *broadcast item*, a *meeting*, etc. In informal or private contexts, a text may be simply a conversation involving a specific group of participants. Alternatively, researchers may elect to define spoken texts solely in terms of their duration in time or length in words. By default, these Guidelines assume of a text only that:

- it is internally cohesive,
- it is describable by a single header, and
- it represents a single stretch of time with no significant discontinuities.

Within a <text> it may be necessary to identify subdivisions of various kinds, if only for convenience of handling. The neutral  element discussed in section *4.1. Divisions of the Body* is recommended for this purpose. It may be found useful also for representing subdivisions relating to discourse structure, speech act theory, transactional analysis, etc., provided only that these divisions are hierarchically well-behaved. Where they are not, as is often the case, the mechanisms discussed in chapters *17. Linking, Segmentation, and Alignment* and *21. Non-hierarchical Structures* may be used.

A spoken text may contain any of the following components:

- utterances
- pauses
- vocalized but non-lexical phenomena such as coughs
- kinesic (non-verbal, non-lexical) phenomena such as gestures
- entirely non-linguistic incidents occurring during and possibly influencing the course of speech
- writing, regarded as a special class of incident in that it can be transcribed, for example captions or overheads displayed during a lecture
- shifts or changes in vocal quality

Elements to represent all of these features of spoken language are discussed in section *8.3. Elements Unique to Spoken Texts* below.

An utterance (tagged <u>) may contain lexical items interspersed with pauses and non-lexical vocal sounds; during an utterance, non-linguistic incidents may occur and written materials may be presented. The <u> element can thus contain any of the other elements listed, interspersed with a transcription of the lexical items of the utterance; the other elements may all appear between utterances or next to each other, but except for <writing> they do not contain any other elements nor any data.

A spoken text itself may be without substructure, that is, it may consist simply of units such as utterances or pauses, not grouped together in any way, or it may be subdivided. If the notion of what constitutes a 'text' in spoken discourse is inevitably rather an arbitrary one, the notion of formal subdivisions within such a 'text' may appear even more debatable. Nevertheless, such divisions may be useful for such types of discourse as debates, broadcasts, etc., where structural subdivisions can easily be identified, or more generally wherever it is desired to aggregate utterances or other parts of a transcript into units smaller than a complete 'text'. Examples might include 'conversations' or 'discourse fragments', or more narrowly, 'that part of the conversation where topic x was discussed', provided only that the set of all such divisions is coextensive with the text.

Each such division of a spoken text should be represented by the numbered or unnumbered  elements defined in chapter *4. Default Text Structure*. For some detailed kinds of analysis a hierarchy of such divisions may be found useful; nested  elements may be used for this purpose, as in the following example showing how a collection made up of transcribed 'sound bites' taken from speeches given by a politician on different occasions might be encoded. Each extract is regarded as a distinct , nested within a single composite  as follows:

```
<div type="soundbites"
  subtype="conservative" org="composite">
  <div sample="medial"/>
  <div sample="medial"/>
  <div sample="initial"/>
</div>
```

As a member of the class `att.declaring`, the  element may also carry a *decls* attribute, for use where the divisions of a text do not all share the same set of the contextual declarations specified in the TEI header. (See further section *16.3. Associating Contextual Information with a Text*).

## 8.2 Documenting the Source of Transcribed Speech

Where a computer file is derived from a spoken text rather than a written one, it will usually be desirable to record additional information about the recording or broadcast which constitutes its source. Several additional elements are provided for this purpose within the source description component of the TEI header:

`<scriptStmt>` (script statement) contains a citation giving details of the script used for a spoken text.

`<recordingStmt>` (recording statement) describes a set of recordings used as the basis for transcription of a spoken text.

`<recording>` (recording event) provides details of an audio or video recording event used as the source of a spoken text, either directly or from a public broadcast.

`@type` the kind of recording.

`<transcriptionDesc>` (transcription description) describes the set of transcription conventions used, particularly for spoken material.

`@ident` supplies an identifier for the encoding convention, independent of any version number.

`@version` supplies a version number for the encoding conventions used, if any.

As a member of the `att.duration` class, the <recording> element inherits the following attribute:

`att.duration.w3c` provides attributes for recording normalized temporal durations.

`@dur` (duration) indicates the length of this element in time.

Note that detailed information about the participants or setting of an interview or other transcript of spoken language should be recorded in the appropriate division of the profile description, discussed in chapter *16. Language Corpora*, rather than as part of the source description. The source description is used to hold information only about the source from which the transcribed speech was taken, for example, any script being read and any technical details of how the recording was produced. If the source was a previously-created transcript, it should be treated in the same way as any other source text.

The <scriptStmt> element should be used where it is known that one or more of the participants in a spoken text is speaking from a previously prepared script. The script itself should be documented in the same way as any other written text, using one of the three citation tags mentioned above. Utterances or groups of utterances may be linked to the script concerned by means of the *decls* attribute, described in section *16.3. Associating Contextual Information with a Text*.

```
<sourceDesc>
  <scriptStmt  xml:id="CNN12">
    <bibl>
      <author>CNN Network News</author>
      <title>News headlines</title>
      <date  when="1991-06-12">12 Jun 91</date>
    </bibl>
  </scriptStmt>
</sourceDesc>
```

The <recordingStmt> is used to group together information relating to the recordings from which the spoken text was transcribed. The element may contain either a prose description or, more helpfully, one or more <recording> elements, each corresponding with a particular recording. The linkage between utterances or groups of utterances and the relevant recording statement is made by means of the *decls* attribute, described in section *16.3. Associating Contextual Information with a Text*.

The <recording> element should be used to provide a description of how and by whom a recording was made. This information may be provided in the form of a prose description, within which such items as statements of responsibility, names, places, and dates may be identified using the appropriate phrase-level tags. Alternatively,

a selection of elements from the `model.recordingPart` class may be provided. This element class makes available the following elements:

`<date>` (date) contains a date in any format.

`<time>` (time) contains a phrase defining a time of day in any format.

`<respStmt>` (statement of responsibility) supplies a statement of responsibility for the intellectual content of a text, edition, recording, or series, where the specialized elements for authors, editors, etc. do not suffice or do not apply. May also be used to encode information about individuals or organizations which have played a role in the production or distribution of a bibliographic work.

`<equipment>` (equipment) provides technical details of the equipment and media used for an audio or video recording used as the source for a spoken text.

`<broadcast>` (broadcast) describes a broadcast used as the source of a spoken text.

Specialized collections may wish to add further sub-elements to these major components. These elements should be used only for information relating to the recording process itself; information about the setting or participants (for example) is recorded elsewhere: see sections *16.2.3. The Setting Description* and *16.2.2. The Participant Description*.

```
<recordingStmt>
  <recording  type="video">
    <p>U-matic recording made by college audio-visual department staff,
        available as PAL-standard VHS transfer or sound-only cassette</p>
  </recording>
</recordingStmt>
```

```
<recordingStmt>
  <recording  type="audio"  dur="P30M">
    <respStmt>
      <resp>Location recording by</resp>
      <name>Sound Services Ltd.</name>
    </respStmt>
    <equipment>
      <p>Multiple close microphones mixed down to stereo Digital
            Audio Tape, standard play, 44.1 KHz sampling frequency</p>
    </equipment>
    <date>12 Jan 1987</date>
  </recording>
</recordingStmt>
```

```
<recordingStmt>
  <recording  type="audio"  dur="P15M"
    xml:id="rec-3001">
    <date>14 Feb 2001</date>
  </recording>
  <recording  type="audio"  dur="P15M"
    xml:id="rec-3002">
    <date>17 Feb 2001</date>
  </recording>
  <recording  type="audio"  dur="P15M"
    xml:id="rec-3003">
    <date>22 Feb 2001</date>
  </recording>
</recordingStmt>
```

When a recording has been made from a public broadcast, details of the broadcast itself should be supplied within the <recording> element, as a nested <broadcast> element. A broadcast is closely analogous to a publication and the <broadcast> element should therefore contain one or the other of the bibliographic citation elements <bibl>, <biblStruct>, or <biblFull>. The broadcasting agency responsible for a broadcast is regarded as its author, while other participants (for example interviewers, interviewees, script writers, directors, producers, etc.) should be specified using the <respStmt> or <editor> element with an appropriate <resp> (see further section *3.12. Bibliographic Citations and References*).

```
<recording type="audio" dur="P10M">
  <equipment>
    <p>Recorded from FM Radio to digital tape</p>
  </equipment>
  <broadcast>
    <bibl>
      <title>Interview on foreign policy</title>
      <author>BBC Radio 5</author>
      <respStmt>
        <resp>interviewer</resp>
        <name>Robin Day</name>
      </respStmt>
      <respStmt>
        <resp>interviewee</resp>
        <name>Margaret Thatcher</name>
      </respStmt>
      <series>
        <title>The World Tonight</title>
      </series>
      <note>First broadcast on <date when="1989-11-27">27 Nov 1989</date>
      </note>
    </bibl>
  </broadcast>
</recording>
```

When a broadcast contains several distinct recordings (for example a compilation), additional <recording> elements may be further nested within the <broadcast> element.

```
<recording dur="P100M">
  <broadcast>
    <recording/>
  </broadcast>
</recording>
```

The <transcriptionDesc> element can be used to document the particular transcription conventions (use of space, punctuation, special characters etc.) used in making the transcription. A number of sets of such conventions have been defined within particular research communities, or by users of particular transcription tools. The attributes *ident* and *version* may be used to refer to such conventions in a machine tractable way, where this is appropriate.

```
<transcriptionDesc ident="HIAT"
  version="2004"/>
```

## 8.3 Elements Unique to Spoken Texts

The following elements characterize spoken texts, transcribed according to these Guidelines:

`<u>` (utterance) contains a stretch of speech usually preceded and followed by silence or by a change of speaker.

`<pause>` (pause) marks a pause either between or within utterances.

`<vocal>` (vocal) marks any vocalized but not necessarily lexical phenomenon, for example voiced pauses, non-lexical backchannels, etc.

`<kinesic>` (kinesic) marks any communicative phenomenon, not necessarily vocalized, for example a gesture, frown, etc.

`<incident>` (incident) marks any phenomenon or occurrence, not necessarily vocalized or communicative, for example incidental noises or other events affecting communication.

`<writing>` (writing) contains a passage of written text revealed to participants in the course of a spoken text.

`<shift>` (shift) marks the point at which some paralinguistic feature of a series of utterances by any one speaker changes.

The <u> element may appear directly within a spoken text, and may contain any of the others; the others may also appear directly (for example, a <vocal> may appear between two utterances) but cannot contain a <u> element. In terms of the basic TEI model, therefore, we regard the <u> element as analogous to a paragraph, and the others as analogous to 'phrase' elements, but with the important difference that they can exist either as siblings or as children of utterances. The class `model.divPart.spoken` provides the <u> element; the class `model.global.spoken` provides the six other elements listed above.

As members of the `att.ascribed` class, all of these elements share the following attributes:

`att.ascribed` provides attributes for elements representing speech or action that can be ascribed to a specific individual.

- `@who` indicates the person, or group of people, to whom the element content is ascribed.

`att.ascribed.directed` provides attributes for elements representing speech or action that can be directed at a group or individual.

- `@toWhom` indicates the person, or group of people, to whom a speech act or action is directed.

As members of the `att.typed`, `att.timed` and `att.duration` classes, all of these elements except <shift> share the following attribute:

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

- `@type` characterizes the element in some sense, using any convenient classification scheme or typology.
- `@subtype` (subtype) provides a sub-categorization of the element, if needed.

`att.timed` provides attributes common to those elements which have a duration in time, expressed either absolutely or by reference to an alignment map.

- `@start` indicates the location within a temporal alignment at which this element begins.
- `@end` indicates the location within a temporal alignment at which this element ends.

`att.duration.w3c` provides attributes for recording normalized temporal durations.

- `@dur` (duration) indicates the length of this element in time.

Each of these elements is further discussed and specified in sections *8.3.1. Utterances* to *8.3.4. Writing*.

We can show the relationship between four of these constituents of speech using the features *eventive*, *communicative*, *anthropophonic* (for sounds produced by the human vocal apparatus), and *lexical*:

| | eventive | communicative | anthropophonic | lexical |
|---|---|---|---|---|
| incident | + | - | - | - |
| kinesic | + | + | - | - |
| vocal | + | + | + | - |
| utterance | + | + | + | + |

The differences are not always clear-cut. Among *incidents* might be included actions like slamming the door, which can certainly be communicative. *Vocals* include coughing and sneezing, which are usually involuntary noises. Equally, the distinction between utterances and vocals is not always clear, although for many analytic purposes it will

be convenient to regard them as distinct. Individual scholars may differ in the way borderlines are drawn and should declare their definitions in the <editorialDecl> element of the header (see *2.3.3. The Editorial Practices Declaration*).

The following short extract exemplifies several of these elements. It is recoded from a text originally transcribed in the CHILDES format.[68] Each utterance is encoded using a <u> element (see section *8.3.1. Utterances*). The speakers are defined using the <listPerson> element discussed in *14.3.2. The Person Element* and each is given a unique identifier also used to identify their speech. Pauses marked by the transcriber are indicated using the <pause> element (see section *8.3.2. Pausing*). Non-verbal vocal effects such as the child's meowing are indicated either with orthographic transcriptions or with the <vocal> element, and entirely non-linguistic but significant incidents such as the sound of the toy cat are represented by the <incident> elements (see section *8.3.3. Vocal, Kinesic, Incident*).

```
<!-- ... in the <particDesc>: --><listPerson>
  <person xml:id="mar">
<!-- ... -->
  </person>
  <person xml:id="ros">
<!-- ... -->
  </person>
  <person xml:id="fat">
<!-- ... -->
  </person>
</listPerson>
<!-- ... in the <text>: -->
<u who="#mar">you
never <pause/> take this cat for show and tell
<pause/> meow meow</u>
<u who="#ros">yeah well I dont want to</u>
<incident>
  <desc>toy cat has bell in tail which continues to make a tinkling sound</desc>
</incident>
<vocal who="#mar">
  <desc>meows</desc>
</vocal>
<u who="#ros">because it is so old</u>
<u who="#mar">how <choice>
    <orig>bout</orig>
    <reg>about</reg>
  </choice>
  <emph>your</emph> cat <pause/>yours is <emph>new</emph>
  <kinesic>
    <desc>shows Father the cat</desc>
  </kinesic>
</u>
<u trans="pause" who="#fat">thats <pause/> darling</u>
<u who="#mar">
  <seg>no <emph>mine</emph> isnt old</seg>
  <seg>mine is just um a little dirty</seg>
</u>
```

This example also uses some elements common to all TEI texts, notably the <reg> tag for editorial regularization. Unusually stressed syllables have been encoded with the <emph> element. The <seg> element has also been used to segment the last utterance. Further discussion of all of such options is provided in section *8.4. Elements Defined Elsewhere*.

Contextual information is of particular importance in spoken texts, and should be provided by the TEI header of a text. In general, all of the information in a header is understood to be relevant to the whole of the associated text.

[68] The original is a conversation between two children and their parents, recorded in 1987, and discussed in **?**

The element <u> as a member of the att.declaring class, may however specify a different context by means of the *decls* attribute (see further section *16.3. Associating Contextual Information with a Text*).

### 8.3.1 Utterances

Each distinct *utterance* in a spoken text is represented by a <u> element, described as follows:

**<u>** (utterance) contains a stretch of speech usually preceded and followed by silence or by a change of speaker.

**@*trans*** (transition) indicates the nature of the transition between this utterance and the previous one.

Use of the *who* attribute to associate the utterance with a particular speaker is recommended but not required. Its use implies as a further requirement that all speakers be identified by a <person> or <personGrp> element, typically in the TEI header (see section *16.2.2. The Participant Description*), but it may also point to another external source of information about the speaker. Where utterances or other parts of the transcription cannot be attributed with confidence to any particular participant or group of participants, the encoder may choose to create <personGrp> elements with *xml:id* attributes such as various or unknown, and perhaps give the root <listPerson> element an *xml:id* value of all, then point to those as appropriate using *who*.

The *trans* attribute is provided as a means of characterizing the transition from one utterance to the next at a simpler level of detail than that provided by the temporal alignment mechanism discussed in section *17.4. Synchronization*. The value specified applies to the transition from the preceding utterance into the utterance bearing the attribute. For example:[69]

```
<u xml:id="ts_a1" who="#a">Have you heard the</u>
<u xml:id="ts_b1" trans="latching" who="#b">the election results? yes</u>
<u xml:id="ts_a2" trans="pause" who="#a">it's a disaster</u>
<u xml:id="ts_b2" trans="overlap" who="#b">it's a miracle</u>
```

In this example, utterance ts_b1 latches on to utterance ts_a1, while there is a marked pause between ts_b1 and ts_a2. ts_b2 and ts_a2 overlap, but by an unspecified amount. For ways of providing a more precise indication of the degree of overlap, see section *8.4.2. Synchronization and Overlap*.

An utterance may contain either running text, or text within which other basic structural elements are nested. Where such nesting occurs, the *who* attribute is considered to be inherited for the elements <pause>, <vocal>, <shift>, and <kinesic>; that is, a pause or shift (etc.) within an utterance is regarded as being produced by that speaker only, while a pause between utterances applies to all speakers.

Occasionally, an utterance may seem to contain other utterances, for example where one speaker interrupts himself, or when another speaker produces a 'back-channel' while they are still speaking. The present version of these Guidelines does not support nesting of one <u> element within another. The transcriber must therefore decide whether such interruptions constitute a change of utterance, or whether other elements may be used. In the case of self-interruption, the <shift> element may be used to show that the speaker has changed the quality of their speech:

```
<u who="#a">Listen to this <shift new="reading"/>The government is
confident, he said, that the current economic problems will be
completely overcome by June<shift new="normal"/> what nonsense</u>
```

Alternatively the <incident> element described in section *8.3.3. Vocal, Kinesic, Incident* might be used, without transcribing the read material:

[69] For the most part, the examples in this chapter use no sentence punctuation except to mark the rising intonation often found in interrogative statements; for further discussion, see section *8.4.3. Regularization of Word Forms*.

```
<u  who="#a">Listen to this
<incident>
   <desc>reads aloud from newspaper</desc>
  </incident> what
nonsense</u>
```

Often, back-channelling is only semi-lexicalized and may therefore be represented using the <vocal> element:

```
<u  who="#a">So what could I have done <vocal  who="#b">
   <desc>tut-tutting</desc>
  </vocal> about it anyway?</u>
```

Where this is not possible, it is simplest to regard the back-channel as a distinct utterance.

### 8.3.2 Pausing

Speakers differ very much in their rhythm and in particular in the amount of time they leave between words. The following element is provided to mark occasions where the transcriber judges that speech has been paused, irrespective of the actual amount of silence:

`<pause>` (pause) marks a pause either between or within utterances.

A pause contained by an utterance applies to the speaker of that utterance. A pause between utterances applies to all speakers. The *type* attribute may be used to categorize the pause, for example as short, medium, or long; alternatively the attribute *dur* may be used to indicate its length more exactly, as in the following example:

```
<u>Okay <pause  dur="PT2M"/>U-m<pause  dur="PT75S"/>the scene opens up
<pause  dur="PT50S"/> with <pause  dur="PT20S"/> um <pause  dur="PT145S"/> you see
a tree okay?</u>
```

If detailed synchronization of pausing with other vocal phenomena is required, the alignment mechanism defined at section *17.4. Synchronization* and discussed informally below should be used. Note that the *trans* attribute mentioned in the previous section may also be used to characterize the degree of pausing between (but not within) utterances.

### 8.3.3 Vocal, Kinesic, Incident

The presence of non-transcribed semi-lexical or non-lexical phenomena either between or within utterances may be indicated with the following three elements.

`<vocal>` (vocal) marks any vocalized but not necessarily lexical phenomenon, for example voiced pauses, non-lexical backchannels, etc.

`<kinesic>` (kinesic) marks any communicative phenomenon, not necessarily vocalized, for example a gesture, frown, etc.

`<incident>` (incident) marks any phenomenon or occurrence, not necessarily vocalized or communicative, for example incidental noises or other events affecting communication.

The *who* attribute should be used to specify the person or group responsible for a <vocal>, <kinesic>, or <incident> which is contained within an utterance, if this differs from that of the enclosing utterance. The attribute must be supplied for a <vocal>, <kinesic>, or <incident> which is not contained within an utterance.

The *iterated* attribute may be used to indicate that the vocal, kinesic, or incident is repeated, for example laughter as opposed to laugh. These should both be distinguished from laughing, where what is being encoded is a shift in voice quality. For this last case, the <shift> element discussed in section *8.3.6. Shifts* should be used.

A child <desc> element may be used to supply a conventional representation for the phenomenon, for example:

**non-lexical** burp, click, cough, exhale, giggle, gulp, inhale, laugh, sneeze, sniff, snort, sob, swallow, throat, yawn

**semi-lexical** ah, aha, aw, eh, ehm, er, erm, hmm, huh, mm, mmhm, oh, ooh, oops, phew, tsk, uh, uh-huh, uh-uh, um, urgh, yup

Researchers may prefer to regard some semi-lexical phenomena as 'words' within the bounds of the <u> element. See further the discussion at section *8.4.3. Regularization of Word Forms* below. As for all basic categories, the definition should be made clear in the <encodingDesc> element of the TEI header.

Some typical examples follow:

```
<u  who="#jan">This is just delicious</u>
<incident>
  <desc>telephone rings</desc>
</incident>
<u  who="#ann">I'll get it</u>
<u  who="#tom">I used to <vocal>
    <desc>cough</desc>
  </vocal> smoke a lot</u>
<u  who="#bob">
  <vocal>
    <desc>sniffs</desc>
  </vocal>He thinks he's tough
</u>
<vocal  who="#ann">
  <desc>snorts</desc>
</vocal>
<!-- ... elsewhere, e.g., in the <particDesc>: -->
<listPerson>
  <person  xml:id="ann">
<!-- ... -->
  </person>
  <person  xml:id="bob">
<!-- ... -->
  </person>
  <person  xml:id="jan">
<!-- ... -->
  </person>
  <person  xml:id="kim">
<!-- ... -->
  </person>
  <person  xml:id="tom">
<!-- ... -->
  </person>
</listPerson>
```

Note that Ann's snorting could equally well be encoded as follows:

```
<u  who="#ann">
  <vocal>
    <desc>snorts</desc>
  </vocal>
</u>
```

The extent to which encoding of incidents or kinesics is included in a transcription will depend entirely on the purpose for which the transcription was made. As elsewhere, this will depend on the particular research agenda and the extent to which their presence is felt to be significant for the interpretation of spoken interactions.

### 8.3.4 Writing

Written text may also be encountered when speech is transcribed, for example in a television broadcast or cinema performance, or where one participant shows written text to another. The <writing> element may be used to distinguish such written elements from the spoken text in which they are embedded.

`<writing>` (writing) contains a passage of written text revealed to participants in the course of a spoken text.

`@gradual` indicates whether the writing is revealed all at once or gradually.

`att.global.source` provides attributes used by elements to point to an external source.

For example, if speaker A in the breakfast table conversation in section *8.3.1. Utterances* above had simply shown the newspaper passage to her interlocutor instead of reading it, the interaction might have been encoded as follows:

```
<u  who="#a">look at this</u>
<writing  who="#a"  type="newspaper"
  gradual="false">Government claims economic problems
<soCalled>over by June</soCalled>
</writing>
<u  who="#a">what nonsense!</u>
```

If the source of the writing being displayed is known, bibliographic information about it may be stored in a <listBibl> within the <sourceDesc> element of the TEI header, and then pointed to using the *source* attribute. For example, in the following imaginary example, a lecturer displays two different versions of the same passage of text:

```
<sourceDesc>
<!-- ...-->
  <bibl  xml:id="FOL1">Shakespeare First Folio text</bibl>
  <bibl  xml:id="FOL2">Shakespeare Second Folio text</bibl>
<!-- ...-->
</sourceDesc>
<!-- ...-->
<u>[...] now compare the punctuation of lines 12 and 14 in these two
versions of page 42...
<writing  source="#FOL1">[...]</writing>
  <writing  source="#FOL2">[...]</writing>
</u>
```

### 8.3.5 Temporal Information

As noted above, utterances, vocals, pauses, kinesics, incidents, and writing elements all inherit attributes providing information about their position in time from the classes `att.timed` and `att.duration`. These attributes can be used to link parts of the transcription very exactly with points on a timeline, or simply to indicate their duration. Note that if *start* and *end* point to <when> elements whose temporal distance from each other is specified in a timeline, then *dur* is ignored.

The <anchor> element (see *17.5. Correspondence and Alignment*) may be used as an alternative means of aligning the start and end of timed elements, and is required when the temporal alignment involves points within an element.

For further discussion of temporal alignment and synchronization see *8.4.2. Synchronization and Overlap* below.

### 8.3.6 Shifts

A common requirement in transcribing spoken language is to mark positions at which a variety of prosodic features change. Many paralinguistic features (pitch, prominence, loudness, etc.) characterize stretches of speech which are not co-extensive with utterances or any of the other units discussed so far. One simple method of encoding such units is simply to mark their boundaries. An empty element called <shift> is provided for this purpose.

`<shift>` (shift) marks the point at which some paralinguistic feature of a series of utterances by any one speaker changes.

`@feature` a paralinguistic feature.

`@new` specifies the new state of the paralinguistic feature specified.

A <shift> element may appear within an utterance or a segment to mark a significant change in the particular feature defined by its attributes, which is then understood to apply to all subsequent utterances for the same speaker, unless

changed by a new shift for the same feature in the same speaker. Intervening utterances by other speakers do not normally carry the same feature. For example:

```
<u>
  <shift feature="loud" new="f"/>Elizabeth
</u>
<u>Yes</u>
<u>
  <shift feature="loud" new="normal"/>Come and try this <pause/>
  <shift feature="loud" new="ff"/>come on
</u>
```

In this example, the word *Elizabeth* is spoken loudly, the words *Yes* and *Come and try this* with normal volume, and the words *come on* very loudly.

The values proposed here for the *feature* attribute are based on those used by the Survey of English Usage (see further **??**oase 1990); this list may be revised or supplemented using the methods outlined in section *24.3. Customization*.

The *new* attribute specifies the new state of the feature following the shift. If this attribute has the special value normal, the implication is that the feature concerned ceases to be remarkable at this point.

A list of suggested values for each of the features proposed follows:

- tempo

  **a** allegro (fast)

  **aa** very fast

  **acc** accelerando (getting faster)

  **l** lento (slow)

  **ll** very slow

  **rall** rallentando (getting slower)

- loud (for loudness):

  **f** forte (loud)

  **ff** very loud

  **cresc** crescendo (getting louder)

  **p** piano (soft)

  **pp** very soft

  **dimin** diminuendo (getting softer)

- pitch (for pitch range):

  **high** high pitch-range

  **low** low pitch-range

  **wide** wide pitch-range

  **narrow** narrow pitch-range

  **asc** ascending

  **desc** descending

  **monot** monotonous

  **scand** scandent, each succeeding syllable higher than the last, generally ending in a falling tone

- tension:

**sl** slurred

**lax** lax, a little slurred

**ten** tense

**pr** very precise

**st** staccato, every stressed syllable being doubly stressed

**leg** legato, every syllable receiving more or less equal stress

- rhythm:

**rh** beatable rhythm

**arrh** arrhythmic, particularly halting

**spr** spiky rising, with markedly higher unstressed syllables

**spf** spiky falling, with markedly lower unstressed syllables

**glr** glissando rising, like spiky rising but the unstressed syllables, usually several, also rise in pitch relative to each other

**glf** glissando falling, like spiky falling but with the unstressed syllables also falling in pitch relative to each other

- voice (for voice quality):

**whisp** whisper

**breath** breathy

**husk** husky

**creak** creaky

**fals** falsetto

**reson** resonant

**giggle** unvoiced laugh or giggle

**laugh** voiced laugh

**trem** tremulous

**sob** sobbing

**yawn** yawning

**sigh** sighing

A full definition of the sense of the values provided for each feature may be provided either in the encoding description section of the text header (see section *2.3. The Encoding Description*) or as part of a TEI customization, as described in section *24.3. Customization*.

## 8.4 Elements Defined Elsewhere

This section describes the following features characteristic of spoken texts for which elements are defined elsewhere in these Guidelines:

- segmentation below the utterance level
- synchronization and overlap
- regularization of orthography

The elements discussed here are not provided by the module for spoken texts. Some of them are included in the core module and others are contained in the modules for linking and for analysis respectively. The selection of modules and their combination to define a TEI schema is discussed in section *1.2. Defining a TEI Schema*.

### 8.4.1 Segmentation

For some analytic purposes it may be desirable to subdivide the divisions of a spoken text into units smaller than the individual utterance or turn. Segmentation may be performed for a number of different purposes and in terms of a variety of speech phenomena. Common examples include units defined both prosodically (by intonation, pausing, etc.) and syntactically (clauses, phrases, etc.) The term *macrosyntagm* has been used by a number of researchers to define units peculiar to speech transcripts.[70]

These Guidelines propose that such analyses be performed in terms of neutrally-named *segments*, represented by the <seg> element, which is discussed more fully in section *17.3. Blocks, Segments, and Anchors*. This element may take a *type* attribute to specify the kind of segmentation applicable to a particular segment, if more than one is possible in a text. A full definition of the segmentation scheme or schemes used should be provided in the <segmentation> element of the <editorialDecl> element in the TEI header (see *2.3.3. The Editorial Practices Declaration*).

In the first example below, an utterance has been segmented according to a notion of syntactic completeness not necessarily marked by the speech, although in this case a pause has been recorded between the two sentence-like units. In the second, the segments are defined prosodically (an acute accent has been used to mark the position immediately following the syllable bearing the primary accent or stress), and may be thought of as 'tone units'.

```
<u>
  <seg>we went to the pub yesterday</seg>
  <pause/>
  <seg>there was no one there</seg>
</u>
<u>
  <seg>although its an old ide´a</seg>
  <seg>it hasnt been on the mar´ket very long</seg>
</u>
```

In either case, the <segmentation> element in the header of the text should specify the principles adopted to define the segments marked in this way.

When utterances are segmented end-to-end in the same way as the s-units in written texts, the <s> element discussed in chapter *18. Simple Analytic Mechanisms* may be used, either as an alternative or in addition to the more general purpose <seg> element. The <s> element is available without formality in all texts, but does not allow segments to nest within each other.

Where segments of different kinds are to be distinguished within the same stretch of speech, the *type* attribute may be used, as in the following example:

```
<u  who="#T1">
  <seg  type="C">I think </seg>
  <seg  type="C">this chap was writing </seg>
  <seg  type="C">and he <del  type="repeated">said hello</del> said </seg>
  <seg  type="M">hello </seg>
  <seg  type="C">and he said </seg>
  <seg  type="C">I'm going to a gate
     at twenty past seven </seg>
  <seg  type="C">he said </seg>
  <seg  type="M">ok </seg>
  <seg  type="M">right away </seg>
  <seg  type="C">and so <gap  extent="1 syll"/> on they went </seg>
  <seg  type="C">and they were <gap  extent="3 sylls"/>
     writing there </seg>
</u>
```

[70] The term was apparently first proposed by ?, where it is defined as follows: 'A text can be analysed as a sequence of segments which are internally connected by a network of syntactic relations and externally delimited by the absence of such relations with respect to neighbouring segments. Such a segment is a syntactic unit called a macrosyntagm' (trans. S. Johansson).

In this example, recoded from a corpus of language-impaired speech prepared by Fletcher and Garman, the speaker's utterance has been fully segmented into clausal (`type="C"`) or minor (`type="M"`) units.

For some features, it may be more appropriate or convenient to introduce a new element in a custom namespace:

```
<u  who="#T1"
   xmlns:ext="http://www.example.org/ns/nonTEI">
<!-- ... -->
  <seg  type="C">and he said </seg>
  <seg  type="C">I'm going to a
  <ext:paraphasia>gate</ext:paraphasia>
    at twenty past seven </seg>
<!-- ... -->
</u>
```

Here, `<ext:paraphasia>` has been used to define a particular characteristic of this corpus for which no element exists in the TEI scheme. See further chapter *24.3. Customization* for a discussion of the way in which this kind of user-defined extension of the TEI scheme may be performed and chapter *1. The TEI Infrastructure* for the mechanisms on which it depends.

This example also uses the core elements <gap> and <del> to mark editorial decisions concerning matter completely omitted from the transcript (because of inaudibility), and words which have been transcribed but which the transcriber wishes to exclude from the segment because they are repeated, respectively. See section *3.5. Simple Editorial Changes* for a discussion of these and related elements.

It is often the case that the desired segmentation does not respect utterance boundaries; for example, syntactic units may cross utterance boundaries. For a detailed discussion of this problem, and the various methods proposed by these Guidelines for handling it, see chapter *21. Non-hierarchical Structures*. Methods discussed there include these:

- 'milestone' tags may be used; the special-purpose <shift> tag discussed in section *8.3.6. Shifts* is an extension of this method
- where several discontinuous segments are to be grouped together to form a syntactic unit (e.g. a phrasal verb with interposed complement), the <join> element may be used

### 8.4.2 Synchronization and Overlap

A major difference between spoken and written texts is the importance of the temporal dimension to the former. As a very simple example, consider the following, first as it might be represented in a playscript:

```
 Jane: Have you read Vanity Fair?
Stig: Yes
Lou: (nods vigorously)
```

To encode this, we first define the participants:

```
<listPerson>
  <person  xml:id="stig">
<!-- ... -->
  </person>
  <person  xml:id="lou">
<!-- ... -->
  </person>
  <person  xml:id="jane">
<!-- ... -->
  </person>
</listPerson>
```

Let us assume that Stig and Lou respond to Jane's question before she has finished asking it—a fairly normal situation in spontaneous speech. The simplest way of representing this *overlap* would be to use the *trans* attribute previously discussed:

```
<u who="#jane">have you read Vanity Fair</u>
<u trans="overlap" who="#stig">yes</u>
```

However, this does not allow us to indicate either the extent to which Stig's utterance is overlapped, nor does it show that there are in fact three things which are synchronous: the end of Jane's utterance, Stig's whole utterance, and Lou's kinesic. To overcome these problems, more sophisticated techniques, employing the mechanisms for pointing and alignment discussed in detail in section *17.4. Synchronization*, are needed. If the module for linking has been enabled (as described in section *8.4.1. Segmentation* above), one way to represent the simple example above would be as follows:

```
<u xml:id="utt1" who="#jane">have you read Vanity <anchor synch="#utt2
#k1" xml:id="a1"/> Fair</u>
<u xml:id="utt2" who="#stig">yes</u>
<kinesic xml:id="k1" who="#lou"
  iterated="true">
  <desc>nods head vertically</desc>
</kinesic>
```

For a full discussion of this and related mechanisms, section *17.4.2. Placing Synchronous Events in Time* should be consulted. The rest of the present section, which should be read in conjunction with that more detailed discussion, presents a number of ways in which these mechanisms may be applied to the specific problem of representing temporal alignment, synchrony, or overlap in transcribing spoken texts.

In the simple example above, the first utterance (that with identifier utt1) contains an <anchor> element, the function of which is simply to mark a point within it. The *synch* attribute associated with this anchor point specifies the identifiers of the other two elements which are to be synchronized with it: specifically, the second utterance (utt2) and the kinesic (k1). Note that one of these elements has content and the other is empty.

This example demonstrates only a way of indicating a point within one utterance at which it can be synchronized with another utterance and a kinesic. For more complex kinds of alignment, involving possibly multiple synchronization points, an additional element is provided, known as a <timeline>. This consists of a series of <when> elements, each representing a point in time, and bearing attributes which indicate its exact temporal position relative to other elements in the same timeline, in addition to the sequencing implied by its position within it.

For example:

```
<timeline unit="s" origin="#TS-P1">
  <when xml:id="TS-P1"
    absolute="12:20:01+01:00"/>
  <when xml:id="TS-P2" interval="4.5"
    since="#TS-P1"/>
  <when xml:id="TS-P6"/>
  <when xml:id="TS-P3" interval="1.5"
    since="#TS-P6"/>
</timeline>
```

This timeline represents four points in time, named TS-P1, TS-P2, TS-P6, and TS-P3 (as with all attributes named *xml:id* in the TEI scheme, the names must be unique within the document but have no other significance). TS-P1 is located absolutely, at 12:20:01:01 BST. TS-P2 is 4.5 seconds later than TS-P2 (i.e. at 12:20:46). TS-P6 is at some unspecified time later than TS-P2 and previous to TS-P3 (this is implied by its position within the timeline, as no attribute values have been specified for it). The fourth point, TS-P3, is 1.5 seconds later than TS-P6.

One or more such timelines may be specified within a spoken text, to suit the encoder's convenience. If more than one is supplied, the *origin* attribute may be used on each to specify which other <timeline> element it follows. The *unit* attribute indicates the units used for timings given on <when> elements contained by the alignment map. Alternatively, to avoid the need to specify times explicitly, the *interval* attribute may be used to indicate that all the <when> elements in a time line are a fixed distance apart.

Three methods are available for aligning points or elements within a spoken text with the points in time defined by the <timeline>:

- The elements to be synchronized may specify the identifier of a <when> element as the value of one of the *start*, *end*, or *synch* attributes
- The <when> element may specify the identifiers of all the elements to be synchronized with it using the *synch* attribute
- A free-standing <link> element may be used to associate the <when> element and the elements synchronized with it by specifying their identifiers as values for its *target* attribute.

For example, using the timeline given above:

```
<u xml:id="TS-U1" start="#TS-P2"
  end="#TS-P3">This is my <anchor synch="#TS-P6" xml:id="TS-P6A"/> turn</u>
```

The start of utterance TS-U1 is aligned with TS-P2 and its end with TS-P3. The transition between the words *my* and *turn* occurs at point TS-P6A, which is synchronous with point TS-P6 on the timeline.

The synchronization represented by the preceding examples could equally well be represented as follows:

```
<timeline origin="#ts-p1" unit="s">
  <when xml:id="ts-p1"
    absolute="12:20:01+01:00"/>
  <when synch="#ts-u1" xml:id="ts-p2"
    interval="4.5" since="#ts-p1"/>
  <when synch="#ts-x1" xml:id="ts-p6"/>
  <when synch="#ts-u1" xml:id="ts-p3"
    interval="1.5" since="#ts-p6"/>
</timeline>
<u xml:id="ts-u1">This is my <anchor xml:id="ts-x1"/> turn</u>
```

Here, the whole of the object with identifier ts-u1 (the utterance) has been aligned with two different points, ts-p2 and ts-p3. This is interpreted to mean that the utterance spans at least those two points.

Finally, a <linkGrp> may be used as an alternative to the *synch* attribute:

```
<timeline origin="#TS-p1" unit="s">
  <when xml:id="TS-p1" absolute="12:20:01"/>
  <when xml:id="TS-p2" interval="4.5"
    since="#TS-p1"/>
  <when xml:id="TS-p6"/>
  <when xml:id="TS-p3" interval="1.5"
    since="#TS-p6"/>
</timeline>
<u xml:id="TS-u1">
  <anchor xml:id="TS-u1start"/>
This is my <anchor xml:id="TS-x1"/> turn
<anchor xml:id="TS-u1end"/>
</u>
<linkGrp type="synchronous">
  <link target="#TS-u1start #TS-p1"/>
```

```
  <link  target="#TS-u1end #TS-p2"/>
  <link  target="#TS-x1 #TS-p6"/>
</linkGrp>
```

As a further example of the three possibilities, consider the following dialogue, represented first as it might appear in a conventional playscript:

```
Tom: I used to smoke - -
Bob: (interrupting) You used to smoke?
Tom: (at the same time) a lot more than this.  But I never
     inhaled the smoke
```

A commonly used convention might be to transcribe such a passage as follows:

```
 (1) I used to smoke [ a lot more than this ]
(2)                  [ you used to smoke ]
(1) but I never inhaled the smoke
```

Such conventions have the drawback that they are hard to generalize or to extend beyond the very simple case presented here. Their reliance on the accidentals of physical layout may also make them difficult to transport and to process computationally. These Guidelines recommend the following mechanisms to encode this.

Where the whole of one or another utterance is to be synchronized, the *start* and *end* attributes may be used:

```
<u  who="#tom">I used to smoke <anchor  xml:id="TS-p10"/> a lot more than this
<anchor  xml:id="TS-p20"/>but I never inhaled the smoke</u>
<u  start="#TS-p10"  end="#TS-p20"  who="#bob">You used to smoke</u>
```

Note that the second utterance above could equally well be encoded as follows with exactly the same effect:

```
<u  who="#bob">
  <anchor  synch="#TS-p10"/>You used to smoke<anchor  synch="#TS-p20"/>
</u>
```

If synchronization with specific timing information is required, a <timeline> must be included:

```
<timeline  origin="#TS-t01"  unit="s">
  <when  xml:id="TS-t01"  absolute="15:33:01Z"/>
  <when  xml:id="TS-t02"  interval="2.5"
    since="#TS-t01"/>
</timeline>
<u  who="#tom">I used to smoke
<anchor  synch="#TS-t01"/>a lot more than this
<anchor  synch="#TS-t02"/>but I never inhaled the smoke</u>
<u  who="#bob">
  <anchor  synch="#TS-t01"/>You used to smoke<anchor  synch="#TS-t02"/>
</u>
```

(Note that If only the ordering or sequencing of utterances is needed, then specific timing information shown here in *unit*, *absolute* and *interval* does not need to be provided.)

As above, since the whole of Bob's utterance is to be aligned, the *start* and *end* attributes may be used as an alternative to the second pair of <anchor> elements:

```
<u  start="#TS-t01"  end="#TS-t02"  who="#bob">You used to smoke</u>
```

An alternative approach is to mark the synchronization by pointing from the <timeline> to the text:

```
<timeline  origin="#TS-T01">
  <when  synch="#TS-nm1 #bob-u2"
    xml:id="TS-T01"/>
  <when  synch="#TS-nm2 #bob-u2"
    xml:id="TS-T02"/>
</timeline>
<u  who="#tom">I used to smoke
<anchor  xml:id="TS-nm1"/>a lot more than this
<anchor  xml:id="TS-nm2"/>but I never inhaled the smoke</u>
<u  xml:id="bob-u2"  who="#bob">You used to smoke</u>
```

To avoid deciding whether to point from the timeline to the text or vice versa, a <linkGrp> may be used:

```
<body>
  <timeline  origin="#T001">
    <when  xml:id="T001"/>
    <when  xml:id="T002"/>
  </timeline>
  <u  who="#tom">I used to smoke
  <anchor  xml:id="NM01"/>a lot more than this
  <anchor  xml:id="NM02"/>but I never inhaled the smoke</u>
  <u  xml:id="bob-U2"  who="#bob">You used to smoke</u>
  <linkGrp  type="synchronize">
    <link  target="#T001 #NM01 #bob-U2"/>
    <link  target="#T002 #NM02 #bob-U2"/>
  </linkGrp>
</body>
```

Note that in each case, although Bob's utterance follows Tom's sequentially in the text, it is aligned temporally with its middle, without any need to disrupt the normal syntax of the text.

As a final example, consider the following exchange, first as it might be represented using a musical-score-like notation, in which points of synchronization are represented by vertical alignment of the text:

```
 Stig: This is |my  |turn
Jane:         |Balderdash
Lou :         |No, |it's mine
```

All three speakers are simultaneous at the words *my*, *Balderdash*, and *No*; speakers Stig and Lou are simultaneous at the words *turn* and *it's*. This could be encoded as follows, using pointers from the alignment map into the text:

```
<timeline  origin="#TSp1">
  <when  synch="#TSa1 #TSb1 #TSc1"
    xml:id="TSp1"/>
  <when  synch="#TSa2 #TSc2"  xml:id="TSp2"/>
</timeline>
<!-- ... -->
<u  who="#stig">this is <anchor  xml:id="TSa1"/> my <anchor  xml:id="TSa2"/> turn</u>
<u  who="#jane"  xml:id="TSb1">balderdash</u>
<u  who="#lou"  xml:id="TSc1"> no <anchor  xml:id="TSc2"/> it's mine</u>
```

### 8.4.3 Regularization of Word Forms

When speech is transcribed using ordinary orthographic notation, as is customary, some compromise must be made between the sounds produced and conventional orthography. Particularly when dealing with informal, dialectal, or other varieties of language, the transcriber will frequently have to decide whether a particular sound is to be treated as a distinct vocabulary item or not. For example, while in a given project *kinda* may not be worth distinguishing as a vocabulary item from *kind of*, *isn't* may clearly be worth distinguishing from *is not*; for some purposes, the regional variant *isnae* might also be worth distinguishing in the same way.

One rule of thumb might be to allow such variation only where a generally accepted orthographic form exists, for example, in published dictionaries of the language register being encoded; this has the disadvantage that such dictionaries may not exist. Another is to maintain a controlled (but extensible) set of normalized forms for all such words; this has the advantage of enforcing some degree of consistency among different transcribers. Occasionally, as for example when transcribing abbreviations or acronyms, it may be felt necessary to depart from conventional spelling to distinguish between cases where the abbreviation is spelled out letter by letter (e.g. *B B C* or *V A T*) and where it is pronounced as a single word (*VAT* or *RADA*). Similar considerations might apply to pronunciation of foreign words (e.g. *Monsewer* vs. *Monsieur*).

In general, use of punctuation, capitalization, etc., in spoken transcripts should be carefully controlled. It is important to distinguish the transcriber's intuition as to what the punctuation should be from the marking of prosodic features such as pausing, intonation, etc.

Whatever practice is adopted, it is essential that it be clearly and fully documented in the editorial declarations section of the header. It may also be found helpful to include normalized forms of non-conventional spellings within the text, using the elements for simple editorial changes described in section *3.5. Simple Editorial Changes* (see further section *8.4.5. Speech Management*).

### 8.4.4 Prosody

In the absence of conventional punctuation, the marking of prosodic features assumes paramount importance, since these structure and organize the spoken message. Indeed, such prosodic features as points of primary or secondary stress may be represented by specialized punctuation marks, or other characters such as those provided by the Unicode Spacing Modifier Letters block. Pauses have already been dealt with in section *8.3.2. Pausing*; while tone units (or intonational phrases) can be indicated by the segmentation tag discussed in section *8.4.1. Segmentation*. The <shift> element discussed in section *8.3.6. Shifts* may also be used to encode some prosodic features, for example where all that is required is the ability to record shifts in voice quality.

In a more detailed phonological transcript, it is common practice to include a number of conventional signs to mark prosodic features of the surrounding or (more usually) preceding speech. Such signs may be used to record, for example, particular intonation patterns, truncation, vowel quality (long or short) etc. These signs may be preserved in a transcript either by using conventional punctuation or by marking their presence by <g> elements. Where a transcript includes many phonetic or phonemic aspects, it will generally be more convenient to use the appropriate Unicode characters (see further chapters *vi Languages and Character Sets* and *5. Characters, Glyphs, and Writing Modes*). For representation of phonemic information, the use of the International Phonetic Alphabet, which can be represented in Unicode characters, is recommended.

In the following example, special characters have been defined as follows within the <encodingDesc> of the TEI header

```
<charDecl>
  <char  xml:id="lf">
    <desc>low fall intonation</desc>
  </char>
  <char  xml:id="lr">
    <desc>low rise intonation</desc>
  </char>
  <char  xml:id="fr">
    <desc>fall rise intonation</desc>
  </char>
  <char  xml:id="rf">
    <desc>rise fall intonation</desc>
```

```
  </char>
  <char  xml:id="long">
    <desc>lengthened syllable</desc>
  </char>
  <char  xml:id="short">
    <desc>shortened syllable</desc>
  </char>
</charDecl>
```

These declarations might additionally provide information about how the characters concerned should be rendered, their equivalent IPA form, etc. In the transcript itself references to them can then be included as follows:

```
<!-- ... in the <particDesc>: --><listPerson>
  <person  xml:id="cwn">
    <p>Customer WN</p>
  </person>
  <person  xml:id="aj">
    <p>Assistant K</p>
  </person>
</listPerson>
<!-- ... within the <text>: -->
<div  n="Lod E-03"  type="exchange">
  <note>C is with a friend</note>
  <u  who="#cwn">
    <unclear>Excuse me<g  ref="#lf"/>
    </unclear>
    <pause/> You dont have some
     aesthetic<g  ref="#short"/>
    <pause/>
    <unclear>specially on early</unclear>
     aesthetics terminology <g  ref="#lr"/>
  </u>
  <u  who="#aj"> No<g  ref="#lf"/>
    <pause/>No<g  ref="#lf"/>
    <gap  extent="2 beats"/> I'm
     afraid<g  ref="#lf"/>
  </u>
  <u  trans="latching"  who="#cwn"> No<g  ref="#lr"/>
    <unclear>Well</unclear> thanks<g  ref="#lr"/>
    <pause/> Oh<g  ref="#short"/>
    <unclear>you couldnt<g  ref="#short"/> can we</unclear> kind of<g  ref="#long"/>
    <pause/>I mean ask you to order it for us<g  ref="#long"/>
    <g  ref="#fr"/>
  </u>
  <u  trans="latching"  who="#aj"> Yes<g  ref="#fr"/> if you know the
title<g  ref="#lf"/> Yeah<g  ref="#lf"/>
  </u>
  <u  who="#cwn">
    <gap  extent="4 beats"/>
  </u>
  <u  who="#aj"> Yes thats fine. <unclear>just as soon as it comes in we'll send
        you a postcard<g  ref="#lf"/>
    </unclear>
  </u>
</div>
```

This example, which is taken from a corpus of bookshop service encounters, also demonstrates the use of the <unclear> and <gap> elements discussed in section *3.5. Simple Editorial Changes*. Where words are so unclear that only their extent can be recorded, the empty <gap> element may be used; where the encoder can identify the words but wishes to record a degree of uncertainty about their accuracy, the <unclear> element may be used. More flexible and detailed methods of indicating uncertainty are discussed in chapter *22. Certainty, Precision, and Responsibility*.

For more detailed work, involving a detailed phonological transcript including representation of stress and pitch patterns, it is probably best to maintain the prosodic description in parallel with the conventional written transcript, rather than attempt to embed detailed prosodic information within it. The two parallel streams may be aligned with each other and with other streams, for example an acoustic encoding, using the general alignment mechanisms discussed in section *8.3.6. Shifts*.

### 8.4.5 Speech Management

Phenomena of *speech management* include disfluencies such as filled and unfilled pauses, interrupted or repeated words, corrections, and reformulations as well as interactional devices asking for or providing feedback. Depending on the importance attached to such features, transcribers may choose to adopt conventionalized representations for them (as discussed in section *8.4.3. Regularization of Word Forms* above), or to transcribe them using IPA or some other transcription system. To simplify analysis of the lexical features of a speech transcript, it may be felt useful to 'tidy away' many of these disfluencies. Where this policy has been adopted, these Guidelines recommend the use of the tags for simple editorial intervention discussed in section *3.5. Simple Editorial Changes*, to make explicit the extent of regularization or normalization performed by the transcriber.

For example, false starts, repetition, and truncated words might all be included within a transcript, but marked as editorially deleted, in the following way:

```
<u>
  <del  type="truncation">s</del>see
<del  type="repetition">you you</del> you know
<del  type="falseStart">it's</del> he's crazy
</u>
```

As previously noted, the <gap> element may be used to mark points within a transcript where words have been omitted, for example because they are inaudible, as in the following example in which 5 seconds of speech is drowned out by an external event:

```
<gap  reason="passing-truck"  quantity="5"
  unit="s"/>
```

The <unclear> element may be used to mark words which have been included although the transcriber is unsure of their accuracy:

```
<u>...and then <unclear  reason="passing-truck">marbled queen</unclear>
</u>
```

Where a transcriber is believed to have incorrectly identified a word, the elements <corr> or <sic> embedded within a <choice> element may be used to indicate both the original and a corrected form of it:

```
<choice>
  <corr>SCSI</corr>
  <sic>skuzzy</sic>
</choice>
```

These elements are further discussed in section *3.5.1. Apparent Errors*.

Finally phenomena such as *code-switching*, where a speaker switches from one language to another, may easily be represented in a transcript by using the <foreign> element provided by the core tagset:

```
<u  who="#P1">I proposed that <foreign  xml:lang="de"> wir können
  <pause  dur="PT1S"/> vielleicht </foreign> go to warsaw
and <emph>vienna</emph>
</u>
```

### 8.4.6 Analytic Coding

The recommendations made here only concern the establishment of a basic text. Where a more sophisticated analysis is needed, more sophisticated methods of markup will also be appropriate, for example, using stand-off markup to indicate multiple segmentation of the stream of discourse, or complex alignment of several segments within it. Where additional annotations (sometimes called 'codes' or 'tags') are used to represent such features as linguistic word class (noun, verb, etc.), type of speech act (imperative, concessive, etc.), or information status (theme/rheme, given/new, active/semi-active/new), etc., a selection from the general purpose analytic tools discussed in chapters *17. Linking, Segmentation, and Alignment*, *18. Simple Analytic Mechanisms*, and *19. Feature Structures* may be used to advantage.

The general-purpose <annotationBlock> element may be used to group together a transcription and multiple layers of annotation. It also serves to divide a transcribed text up into meaningful analytic sections.

`<annotationBlock>` groups together various annotations, e.g. for parallel interpretations of a spoken segment.

## 8.5 Module for Transcribed Speech

The module described in this chapter makes available the following components:

**Module spoken: Transcribed Speech**

- *Elements defined*: annotationBlock broadcast equipment incident kinesic pause recording recordingStmt scriptStmt shift transcriptionDesc u vocal writing
- *Classes defined*: att.duration model.divPart.spoken model.global.spoken model.recordingPart

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 9 Computer-mediated Communication

This chapter describes the TEI encoding mechanisms available for textual data that represents discourse from genres of computer-mediated communication (CMC). It is intended to provide the basic framework needed to encode CMC corpora.

## 9.1 General Considerations

While the term *computer-mediated communication* might be used broadly to describe all kinds of communications that are mediated by digital technologies (such as text on web pages, written exchanges in chats and forums, interactions with artificial intelligence systems, the spoken conversations in internet video meetings), for the purposes of these Guidelines we use the term to apply to forms of communication that share the following features:

- they are dialogic;
- they are organized as interactional sequences so that each communicative move may determine the context for subsequent moves (typically taken by another interlocutor) and may react to the context created by a prior move;
- they are created and displayed using computer technology or human-machine interfaces such as keyboard, mouse, speech-to-text conversion software, monitor or screen and transmitted over a computer network (typically the internet).

Such communications may be expressed as posts (cf. *9.3.1. CMC Posts*), utterances, onscreen activities, or bodily activities exerted by a virtual avatar.

The following kinds of platforms support CMC:

- chats, messengers, or online forums;
- social media platforms and applications;
- the communication functions of collaborative platforms and projects (e.g. an online learning environment, or a 'talk' page);
- 3D virtual world environments;
- other interactive services supported by the internet.

CMC supports multimodal expression combining text, images, sound. Whereas early CMC systems (e.g. Internet Relay Chat, 'IRC' for short, the Usenet 'newsgroups', or even the Unix talk system) were completely ASCII-based, most CMC applications now permit combining media formats (e.g. written or spoken language with graphic icons and images) and mixing communication technologies on one platform (e.g. combined use of an audio connection, a chat system, and a 3D interface in which users control a virtual avatar as in many multiplayer online computer games or in virtual worlds).

## 9.2 Basic Units of CMC

This section describes the encoding mechanisms for the basic units of CMC and for their combined use to encode CMC data.

We use the term *basic CMC unit* to refer to a communication produced by an interlocutor to initiate or contribute to an ongoing CMC interaction or joint CMC activity. Contributions to an ongoing interaction are produced to perform a move to develop the interactional sequence, for instance to respond in chats or forum discussions. Contributions to joint CMC activities may not all be directly interactional; some may be part of a collaborative project of the involved individuals. Such collaboration could involve editing activities in a shared text editor or whiteboard in parallel with an ongoing CMC interaction (chat, audio conversation, or audio-video conference) in the same CMC environment in which these editing activities are discussed by the participants.

Basic units of CMC can be described according to three criteria:

1. the temporal properties of when these contributions are produced by their creators, transmitted via CMC systems, and made accessible for the recipients;
2. the modality of the unit as a whole, whether verbal or nonverbal;
3. for verbal units: whether the unit is expressed in the written or spoken mode.

A taxonomy of basic CMC units resulting from these criteria is given in the following figure.

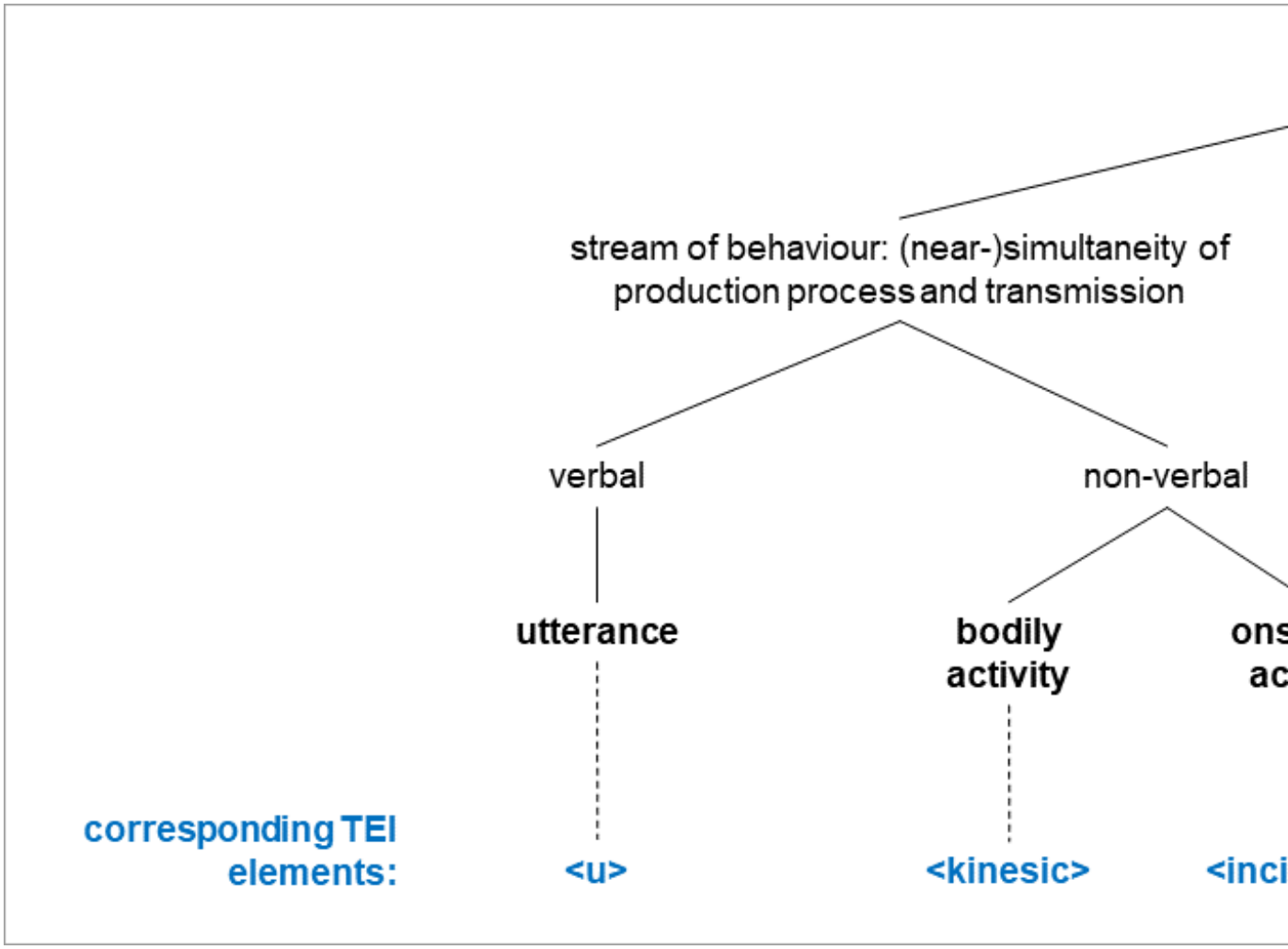



Figure 12: Taxonomy of basic CMC units according to ?

The most important distinction in the CMC taxonomy concerns the temporal nature of units exchanged via CMC technologies. The left part of the taxonomy describes units that are performed (by a producer) and perceived (by a recipient) as a continuous stream of behaviour. Units of this type can be performed as

**spoken utterances**, i.e. stretches of speech which are produced to perform a speaker turn in a conversation,

**bodily activity**, i.e. nonverbal behaviour (gesture, gaze) produced to perform a speaker turn, either performed by the real body of an interlocutor (e.g. in a video conference) or through the virtual avatar of an interlocutor in a 3D environment,

**onscreen activities**, i.e. non-bodily expressions that are transmitted to the group of interacting or coworking participants, for instance the editing of content in a shared text editor which can be perceived by the other parties simultaneously (as may be the case in learning or collaboration environments).

The right part of the CMC taxonomy describes units in which the production, transmission, and perception of contributions to CMC interactions are organized in a strictly consecutive order: The content—verbal, nonverbal, or multimodal—of the contribution has to be produced before it can be transmitted through a network and made available on the computer monitor or mobile screen of any other party as a preserved and persistent unit. We term this type of unit a *post*. Posts occur in two different variants:

- as *written or multimodal posts,* which are produced with an editor form that is designed for the composition of stretches of written text. Most contemporary post-based CMC technologies provide features for the inclusion of graphic and audio-visual content (emoji graphics, images, videos) into posts and even to produce posts without verbal content (which then may consist only of emojis, an image, or a video file). Written and multimodal posts are the standard formats for user contributions in primarily text-based CMC genres and applications such as chat, SMS, WhatsApp, Instagram, Facebook, X (Twitter), online forums, or Wikipedia talk pages.
- as *audio posts*, which are produced using a recording function. In contrast to CMC units of the type *utterance* which are produced and transmitted simultaneously, audio posts first have to be recorded as a whole and are then transmitted, as audio files, via the internet; the availability of the recording is indicated in the screen protocol by a template-generated, visual post; the recipients can play the recording (repeatedly) by activating the play button displayed in the post on the screen. Examples of CMC applications that implement audio posts are WhatsApp or RocketChat.

Three of the four basic CMC units described above can be represented with models that are described elsewhere in the TEI Guidelines:

| CMC unit | Type of corpus data | TEI P5 element |
|---|---|---|
| spoken utterance | transcription of speech | <u> |
| bodily activity | textual description | <kinesic> |
| onscreen activity | textual description | <incident> |

The <u>, <kinesic>, and <incident> elements are not limited to CMC, but apply to encode textual transcriptions of spoken turns and CMC data about bodily activity and onscreen activity. The CMC unit *post*, which is specific to CMC, is introduced in *9.3.1. CMC Posts*.

## 9.3 Encoding Unique to CMC

This section describes elements, attributes, and models which are unique to CMC and the TEI CMC module.

### 9.3.1 CMC Posts

While the concept of a *post* is not unique to computer-mediated communication (ask anyone who has posted a 'lost cat' sign in the local market), this chapter concerns itself only with postings within a framework of a CMC system. Thus the element <post> is unique to the encoding of computer-mediated communication (CMC).

`<post>` a written (or spoken) contribution to a CMC interaction which has been composed (or recorded) by its author in its entirety before being transmitted over a network (e.g., the internet) and made available on the monitor or screen of the other parties en bloc.

Posts occur in a broad range of written CMC genres, including (but not limited to) messages in chats and WhatsApp dialogues, tweets in X (Twitter) timelines, comments on Facebook pages, posts in forum threads, and comments or contributions to discussions on Wikipedia talk pages or in the comment sections of weblogs.

Posts can be either written or spoken:

- *written* or *multimodal posts*: In the majority of CMC technologies posts are composed as stretches of text using a keyboard or speech-to-text conversion software in an entry form on the screen. In many cases the technology allows authors to include or embed graphics (emojis or images), video files, and hyperlinks into their posts.
- *spoken (audio posts)*: A growing number of CMC technologies, e.g. messenger software such as WhatsApp or RocketChat, allow for an alternative, spoken production of posts by providing a recording function which allows users to record a stretch of spoken language and transmit the resulting audio file to the other parties.

The element <post> may co-occur with <u>, <kinesic>, <incident>, or other existing TEI elements within a , or directly within the <body>, and may contain headings, paragraphs, openers, closers, or salutations.

The <post> element is a member of several TEI attribute classes, including `att.ascribed`, `att.canonical`, `att.datable`, `att.global`, `att.timed`, and `att.typed`, and as such may take a variety of attributes.

### 9.3.2 Attributes Specific to CMC <post>

Three attributes pertain specifically to <post>:

`<post>` a written (or spoken) contribution to a CMC interaction which has been composed (or recorded) by its author in its entirety before being transmitted over a network (e.g., the internet) and made available on the monitor or screen of the other parties en bloc.

`@modality` written or spoken mode.

`@replyTo` indicates to which previous post the current post replies or refers.

`att.indentation` provides attributes for describing the indentation of a textual element on the source page or object.

`@indentLevel` specifies the level of indentation of an item using a numeric value.

The type of the content of a post (i.e., whether the content is text, an image, a video clip, etc.) is indicated by the child elements of the <post>. (E.g., a <post> might have a child <p>, or a child <figure> with a <graphic>, or a child <figure> with a <media>, or some combination thereof.) How that content was created—whether it was recorded speech or not—may be described with the *modality* attribute. Because spoken language differs significantly from written language, the suggested values only separate the spoken modality from the written modality, which covers all cases other than spoken natural language. The use of *modality* is recommended but not required.

```
<post  modality="written"
  generatedBy="human"  synch="#t005"  who="#A06"
  xml:id="cmc_post09">
  <figure  type="image"  generatedBy="human">
    <desc  xml:lang="en">screenshot of the google search for hairdresser "Pasha's
Haare'm"
        with the average google rating (4,5 of 5 stars), the address, the phone number,
and
        the opening hours.</desc>
  </figure>
</post>
```

The *replyTo* attribute is used to capture information drawn from the original metadata associated with a post that asserts to which previous post the current post is a response, or to which previous post it refers. This metadata is included by many, but not all, CMC environments, when the user executes a formal reply action (e.g., by clicking or tapping a reply button). This attribute should not be used to encode interpreted or inferred reply relations based on linguistic cues or discourse markers.

The *replyTo* attribute indicates the replied-to or referred-to posts by providing one or more pointers to them. In the following example, reply references in the source indicate that the first <post> is a reply to an initial post that is not part of the example, the second is a reply to the first, and the third is a reply to the second.

```
<post type="comment" modality="written"
  generatedBy="human" xml:id="cmc_post10" who="#u7"
  replyTo="#cmc_post09" when-iso="2015-07-29T21:44">
  <p>Es hat den Anschein, als wäre bei BER durchaus große Kompetenz am Bau, allerdings
     nicht in Form von Handwerkern….</p>
  <p>http://www.zeit.de/2015/29/imtech-flughafen-berlin-ber-verzoegerung/komplettansicht</p>
</post>
<post type="comment" modality="written"
  generatedBy="human" xml:id="cmc_post11" who="#u8"
  replyTo="#cmc_post10" when-iso="2015-07-30T19:11">
  <p>Nein Nein, an den Handwerkern kann es rein strukturel nicht gelegen haben. Niemand
     lässt seine Handwerker auf der Baustelle derart allein. Zudem gibt es
höchstoffizielle
     “Abnahmen” von Bauabschnitten/phasen. Welcher Mangel auch bestanden hatte, er
hätte
     Zeitnah auffallen müssen.</p>
  <p>Uuups, für Imtek hab ich mal in einer Nachunternehmerfirma gearbeitet. Imtek is
     offenbar ein universeler Bauträger, der alles baut.</p>
</post>
<post type="comment" modality="written"
  generatedBy="human" xml:id="cmc_post12" who="#u8"
  replyTo="#cmc_post11" when-iso="2015-07-30T19:26">
  <p>Stahlkunstruktionen dacht ich mal, was die bauen—oder bauen lassen.</p>
  <p>Das ist schon ein übles Ding. Die Ausschreibungenund Angebote sind unauffällig,
aber
     wenn Unregelmässigkeiten auftreten (im Bauverlauf) dann gibt es die saftigen
     Rechnungen. Da steht dann der Bauherr da und fragt sich, wie er denn so schnell
einen
     fähigen Ersatz herbekommt. Und diese Frage erübrigt sich meist, weil der Markt der
     Baufirmen das nicht hergibt – weil tendenziel 100 % Auslastung. (und noch
schlimmer:
     Absprachen) Was auch Folge des Marktdrucks gewesen war.</p>
</post>
```

In the CMC genre of wiki talk, users insert their contribution to a discussion by modifying the wiki page of the discussion—the talk page. Since there is no technical reply action available in wiki software, users apply textual indentation in the wiki code to indicate a reply to a previous message, and a threaded structure is formed by a series of such indentations. The attribute *indentLevel* records the level of indentation, that is the nesting depth of the current post in such a thread-like structure (as defined by its author and in relation to the standard level of non-indentation which should be encoded with an *indentLevel* of 0). It is used in wiki talk corpora but may also be used for other threaded genres, for example when HTML is used as a source.

The following is a sample encoding of a portion of a discussion among four different users on a Wikipedia talk page.

```
<div type="thread" xml:id="i.10031_19">
  <head>[[WP:AUTO]]</head>
  <post indentLevel="0" modality="written"
    when-iso="2006-09-07T03:09+00" who="#WU00010808" xml:id="cmc_post13">
    <p> I would kindly request from Mr. Meyer to allow others to edit the [...]</p>
  </post>
  <post indentLevel="1" modality="written"
    when-iso="2006-09-08T03:49+00" who="#WU00010804" xml:id="cmc_post14">
    <p>I dont agree, this article is not about Dr. Meyer, [...]</p>
  </post>
  <post indentLevel="2" modality="written"
```

```
    when-iso="2006-09-08T04:16+00"  who="#WU00005520"  xml:id="cmc_post15">
    <p>Why don't you read the policy. [...]</p>
  </post>
  <post  indentLevel="3"  modality="written"
    when-iso="2006-11-01T22:58+00"  who="#WU00010815"  xml:id="cmc_post16">
    <p>Because the policy makes no sense, [...]</p>
  </post>
</div>
```

### 9.3.3 Attributes for General CMC Encoding

The attribute *generatedBy* is also unique to CMC encoding. But unlike *modality*, *replyTo*, and *indentLevel*, *generatedBy* is available not only on the <post> element, but on any of its descendants as well.

`att.cmc` (computer-mediated communication) provides attributes categorizing how the element content was created in a CMC environment.

`@generatedBy` (generated by) categorizes how the content of an element was generated in a CMC environment.

The *generatedBy* attribute may indicate, for <post> or any of its descendants, how the content transcribed in an element was generated in a CMC environment. That is, whether the source text being transcribed was created by a human user, created by the CMC system at the request of a human user (e.g., when the user activates a template that generates the content, such as in a signature), generated by the CMC system (e.g. a status message or a timestamp), or generated by an automated process external to the CMC system itself. This attribute is optional; when it is not specified on a <post> element its value is presumed to be unspecified; when it is unspecified on any descendant of <post> its value is inherited from the immediately enclosing element. In turn, if *generatedBy* is not specified on that element it inherits the value from its immediately enclosing element, and so on up the document hierarchy until a <post> is reached; the <post> either has a *generatedBy* attribute specified or its presumed value is unspecified.

A list of suggested values for *generatedBy* follows:

**human** when the content of the respective element was 'naturally' typed or spoken by a human user (cf. the chat posts in example haircut)

**template** when the content of the respective element was generated after a human user activated a template for its insertion (often applicable to <signed> and <time>; e.g. see the signature in wiki talk in this example below)

**system** when the content of the respective element was generated by the system, i.e. the CMC environment (see, e.g., the system message in an IRC chat in the this other example below)

**bot** when the content of the respective element was generated by a bot, i.e. a non-human agent, typically one that is not part of the CMC environment itself

**unspecified** when it is unspecified or unknown how the content of the respective element was generated (see, e.g. the retweet that forms the second <post> in this example below).

The following is a sample encoding of a chat post that contains an emoji. Although the post was written by a human, the emoji itself was marked in the source as having been generated by a template:

```
<post  modality="written"
  generatedBy="human"  synch="#t003"  who="#A02"
  xml:id="cmc_post18"  xml:lang="de"> Da kostet ein Haarschnitt 50 €
<figure  type="emoji"
    generatedBy="template">
    <desc  type="label"  xml:lang="en">face screaming in fear</desc>
    <desc  type="unicode">U+1F631</desc>
  </figure>
</post>
```

In the following example, the user signature of a wiki talk post was inserted by activating a template, and is thus marked accordingly:

```
<post modality="written"
  xml:id="cmc_post19" indentLevel="0" who="#u005" synch="#t005">
  <p>I'm not sure that this is a proper criterium, or even what this means. What if we
set
    an explosion that breaks a comet into two pieces? What if we build a moon? Cheers,
  </p>
  <signed generatedBy="template"
    rend="inline">
    <ref target="/wiki/User:Greenodd">Greenodd</ref> (<ref target="/wiki/User_talk:Greenodd">tal
<time>01:00, 21
        July 2011 (UTC)</time>
  </signed>
</post>
```

In the following example, a tweet is specified as having been written by a human; however inside the tweet, the timestamp is marked as generated by the CMC system:

```
<post modality="written" type="tweet"
  generatedBy="human" synch="#tweetsbcrn18.t001"
  xml:id="cmc_post_1043764753502486528" who="#u1" xml:lang="de">
  <time generatedBy="system"> 12:31 </time> Heute mit super Unterstützung, wir
grunzen,
wenn die Zeit vorbei ist. <ref type="hashtag"
    target="https://twitter.com/hashtag/bcrn18?src=hash">#bcrn18</ref>
  <ref type="hashtag"
    target="https://twitter.com/hashtag/wikidach?src=hash">#wikidach</ref> PS: Die
beiden brauchen noch Namen. Hinweise dazu am Empfang abgeben!
<ref type="twitter-account"
    target="https://twitter.com/AndreLo79">@AndreLo79</ref>
  <figure type="image">
    <graphic url="https://pbs.twimg.com/media/DnwygdSW4AAoTUn.jpg:large"/>
  </figure>
</post>
<post modality="written"
  generatedBy="unspecified" type="tweet" who="#u1"
  synch="#tweetsbcrn18.t002" xml:id="cmc_post_1043769240136880128">
  <ptr type="retweet"
    target="#cmc_post_1043767827927388160"/>
</post>
<post modality="written"
  generatedBy="human" type="tweet" who="#u3"
  synch="#tweetsbcrn18.t002" xml:lang="de"
  xml:id="cmc_post_1043767827927388160">
  <time generatedBy="system"> 12:43 </time>
  <figure type="image" generatedBy="human">
    <graphic url="https://pbs.twimg.com/media/Dnw1TRNXgAAKqlK.jpg:large"/>
  </figure>
</post>
```

Finally, in the following example of an IRC post, the status message that user 'Interseb has entered the room' was generated by the system, i.e. the chat environment.

```
<post type="event" generatedBy="system"
  who="#SYSTEM" rend="color:navy">
```

```
  <p>
    <name  type="nickname"  ref="#A07">Interseb</name> betritt den Raum.</p>
</post>
```

## 9.4 CMC Macrostructure

In many CMC genres, posts may occur in a variety of ways: e.g. in a sequence or in threads, or grouped in some other way. For example, in chat communication such as WhatsApp, posts are part of 'a chat' of one user with another user or among a group of users. When an entire chat is saved, typically a 'logfile' of the chat is obtained from the CMC system and downloaded. Similarly, Wikipedia discussions occur on a *talk page*, which ultimately is a web page containing the user posts, sub-structured in threads. Likewise, YouTube comments occur on a webpage containing the YouTube video along with comment posts and posts replying to those comments displayed below the video. The video serves as a *prompt* for the whole discussion. In forum discussions, the prompt may be a news item, and in Wikipedia, an article may be viewed as the prompt for the discussion on the talk page associated with that article.

### 9.4.1 Macrostructure of CMC Collections and Documents

When CMC documents are compiled into a collection, dataset, or corpus, we distinguish the following levels in the macrostructure of CMC in TEI:

**The corpus level** The level of a corpus or collection of CMC texts of a particular genre, generally obtained from a particular CMC platform, sometimes even from several platforms. This level may be represented by either a <TEI> element or a <teiCorpus> element. The <teiHeader> of the corpus (i.e., the <teiHeader> that is a child of the outermost <TEI> or <teiCorpus>) will contain metadata in its <sourceDesc> about the CMC platform(s). Metadata about the project responsible for collecting the data and building the corpus, if applicable, should be recorded as well.

**The document level** A set of posts collected (or sampled) by a researcher for analysis. The posts of the document will often map directly to the set of posts grouped on an existing web page, thread, or document within a CMC environment. Within the CMC environment the document as such is often created by a particular user, thereby initiating the communication which other users may read, and to which some other users might contribute. This level will naturally be represented by the <TEI> element. The <teiCorpus> (or <TEI>) element that represents the corpus will contain one or more <TEI> elements as usual.

In the <teiHeader> of a document level <TEI>, the <sourceDesc> will contain metadata about the CMC document such as a title, its author or owner, its URL, the date of its creation, the date of the last change made to it, and other metadata that are available and to be recorded such as one or more categories associated with the document.

The document sometimes contains, or is associated with, a prompt such as a video or a news item, either provided by the initiating user herself or located elsewhere and referenced at the beginning of the document. In such cases, the <teiHeader> of the document should also contain metadata about this prompt.

**The post level** The level of the individual post is naturally represented by the <post> element; its encoding is further described in section *9.3.1. CMC Posts*. A <TEI> element will contain a number of <post> elements, which can be grouped or ordered in  elements representing sequences or threads (section *9.4.2. Sequences, Sections, Threads*) if appropriate.

```
<teiCorpus xmlns="http://www.tei-c.org/ns/1.0">
<!-- a corpus, collection or dataset of CMC documents -->
  <teiHeader>
<!-- metadata pertaining to the corpus or CMC dataset-->
  </teiHeader>
  <TEI>
<!-- a CMC document such as a chat log or a discussion page -->
    <teiHeader>
```

```
<!-- metadata pertaining to the CMC document -->
   </teiHeader>
   <text>
     <body>
       <div>
<!-- subdivisions of the CMC document e.g. in sections or threads if applicable-->
         <post>
<!-- one post -->
         </post>
<!-- more posts -->
       </div>
     </body>
   </text>
  </TEI>
<!-- more documents -->
</teiCorpus>
```

### 9.4.2 Sequences, Sections, Threads

As shown in Example *9.3.2. Attributes Specific to CMC post* above, nested threads of posts may be encoded sequentially, while the *indentLevel* attribute of <post> is used to keep track of the original nesting depth. This is especially meant for CMC text obtained from a wiki code or HTML source, where it is not always entirely clear whether the indentation information actually reflects a reply action from a user.

In genres where technical reply information is available for each post, reply links can be encoded using the *replyTo* attribute on <post> elements, as shown in the second example of *9.3.2. Attributes Specific to CMC post.* The network of all reply links will then also form a threaded structure, and visual indentations can be reconstructed from it and need not be explicitly encoded.

Threads may also be explicitly encoded as nested  elements as in the following skeleton:

```
<div type="thread" n="0">
  <post>...</post>
  <div type="thread" n="1">
    <post> ... </post>
    <post> ... </post>
    <div type="thread" n="2">
<!-- posts -->
    </div>
  </div>
</div>
```

Using this encoding strategy, this example from *9.3.2. Attributes Specific to CMC post* could be encoded as follows:

```
<div type="thread" n="0">
  <post type="comment" xml:id="cmc_post01"
    who="#u7" when-iso="2015-07-29T21:44">
    <p>Es hat den Anschein, ...</p>
  </post>
  <div type="thread" n="1">
    <post type="comment" xml:id="cmc_post02"
      who="#u8" when-iso="2015-07-30T19:11">
      <p>Nein Nein, an den Handwerkern kann es ...</p>
    </post>
    <div type="thread" n="2">
      <post type="comment" xml:id="cmc_post03"
```

```
      who="#u8"  when-iso="2015-07-30T19:26">
      <p>Stahlkunstruktionen dacht ich mal, ....</p>
    </post>
  </div>
 </div>
</div>
```

### 9.4.3 Multimodal CMC

As explained in section *9.2. Basic Units of CMC*, the elements <post>, <u>, <kinesic>, and <incident> are available to encode textual transcriptions of written posts, spoken turns, bodily activities of avatars, and onscreen activity by users that occur in CMC data; and, as discussed in section *9.3.2. Attributes Specific to CMC post*, graphics or other media data within posts are encoded in a <post> with *modality* set to written. When two or more of these features occur in a CMC interaction, we can speak of *multimodal* CMC.

Some basic multimodality is available in many private chat systems such as WhatsApp, where spoken and written posts and media posts containing images or video clips can alternate in the sequence of posts. The following shows the suggested encoding of an extended part of the *haircut* chat example from above, including a spoken post, several written posts, and a post containing a graphic image (adapted from the MoCoDa2 corpus ?)

```
<post  modality="spoken"  generatedBy="human"
  synch="#cmc-haircut_t004"  who="#cmc-haircut_A05"
  xml:id="cmc-haircut_m9"> In Düsseldorf gibt's da so Abstufungen. Da gibt's einmal
Oliver
Schmidt, Oliver Schmidt's Hair Design, also dann, ist eher also, keine Ahnung, zum
Beispiel ich war da bei dem etwas Günstigeren dann. Ich weiß nicht, ob's das in Essen
auch
gibt diese Abstufungen </post>
<post  modality="written"
  generatedBy="human"  synch="#cmc-haircut_t004"
  who="#cmc-haircut_A02"  xml:id="cmc-haircut_m10"> Ich schau mal :) </post>
<post  modality="written"
  generatedBy="human"  synch="#cmc-haircut_t005"
  who="#cmc-haircut_A06"  xml:id="cmc-haircut_m11"> Ich gehe immer nach Katernberg zu
Pasha’s
haarem Hahaha also die sind echt entspannt und gut und nicht teuer </post>
<post  modality="written"
  generatedBy="human"  synch="#tcmc-haircut_005"
  who="#cmc-haircut_A06"  xml:id="cmc-haircut_m12">
  <figure  type="image"  generatedBy="human">
    <desc  xml:lang="en">screenshot of the google search for hairdresser "Pasha's
Haare'm"
        with the average google rating (4,5 of 5 stars), the address, the phone number,
and
        the opening hours. </desc>
  </figure>
</post>
<post  modality="written"
  generatedBy="human"  synch="#cmc-haircut_t006"
  who="#cmc-haircut_A03"  xml:id="cmc-haircut_m13"> Olivers hair und Oliver Schmidt
gehören
zusammen </post>
```

In the graphical user interface (GUI) of a more complex multimodal CMC environment such as Second Life, a gaming and learning platform, interactions may consist of interleaved occurrences of posts (<p>), utterances (<u>) and nonverbal acts such as bodily activities (<kinesic>) or other on-screen activities (<incident>). In the following

example a spoken utterance, an avatar's bodily activity, and a written post occur on the same level within the <body> element, representing parts of a multimodal chat in Second Life (adapted from the Archi21 corpus).

```
<text>
  <body>
    <u  xml:id="cmr-archi21-slrefl-es-j3-1-a191"
      who="#tingrabu"
      start="#cmr-archi21-slrefl-es-j3-1-ts373"  end="#cmr-archi21-slrefl-es-j3-1-ts430">ok
         hm for me this presentation was hm <pause  dur="PT1S"/> become too fast because
it's
         always the same in our architecture school euh we have not time and hm
<pause  dur="PT1S"/> too quickly sorry [...]</u>
    <kinesic  xml:id="cmr-archi21-slrefl-es-j3-1-a192"
      who="#romeorez"
      start="#cmr-archi21-slrefl-es-j3-1-ts376"  end="#cmr-archi21-slrefl-es-j3-1-ts377"
      type="body"  subtype="kinesics">
      <desc>
        <code>eat(popcorn)</code>
      </desc>
    </kinesic>
<!-- more bodily activities of avatars -->
    <post  modality="written"
      generatedBy="human"  xml:id="cmr-archi21-slrefl-es-j3-1-a195"
      who="#tfrez2"
      start="#cmr-archi21-slrefl-es-j3-1-ts380"  end="#cmr-archi21-slrefl-es-j3-1-ts381"
      type="chat-message">
      <p>it went too quickly?</p>
    </post>
  </body>
</text>
```

Note that the spoken utterance <u> represents a speaker turn that was transmitted via an audio channel of the application that is continuously open during a session, whereas a spoken <post> represents a spoken message that has been recorded in private and been posted to the CMC server as a whole. See section *9.2. Basic Units of CMC*.

## 9.5 Documenting CMC (and providing general metadata)

### 9.5.1 Documenting the Source of a Corpus of CMC data

The <teiHeader> of the corpus should contain metadata about the CMC platform(s), e.g. its name, information about its owner (often a company) including their address or location, the URL of the server where the CMC data were collected from, or the filename of a database dump that was used as a source. Metadata about the project responsible for collecting the data and building the corpus, if applicable, should be recorded as well.

The following example shows the <sourceDesc> of a X (Twitter) corpus.

```
<sourceDesc>
  <biblFull>
    <titleStmt>
      <title>Twitter Sample</title>
    </titleStmt>
    <publicationStmt>
      <distributor>Twitter International Company</distributor>
      <address>
        <addrLine>1 Cumberland Place</addrLine>
        <addrLine>Fenian Street</addrLine>
        <addrLine>Dublin 2</addrLine>
```

```
      <postCode>D02 AX07</postCode>
      <country>Ireland</country>
    </address>
    <ptr  target="https://twitter.com/"/>
    <date  when="2024-04-27"/>
  </publicationStmt>
 </biblFull>
</sourceDesc>
```

The following example shows how a Wikipedia database dump may be encoded as the source.

```
<sourceDesc>
 <biblFull>
  <titleStmt>
   <title>German Wikipedia Data Dump of 2019-08-01</title>
  </titleStmt>
  <editionStmt>
   <edition>Dump file in XML (compressed)</edition>
  </editionStmt>
  <extent>
   <measure  unit="GiB"  quantity="7.9"/>
  </extent>
  <publicationStmt>
   <publisher>Wikimedia Foundation, Inc.</publisher>
   <pubPlace>
    <ptr  target="https://dumps.wikimedia.org/"/>
   </pubPlace>
   <date  when="2019-08-01">01 Aug 19</date>
   <idno  type="dump-filename">dewiki-2019-08-01-pages-meta-current</idno>
  </publicationStmt>
 </biblFull>
</sourceDesc>
```

### 9.5.2 Describing the Source of a CMC Document

A CMC document may be a chat logfile, a discussion page, or a thematical thread of posts as encoded within a <TEI> element. Among the metadata to be recorded in the <sourceDesc> of its <teiHeader> are, if available, its title, author or owner, its URL, the date of its creation and/or the date of its last change (i.e. the time when the last post was added to it).

The following example is the <sourceDesc> of a TEI encoding of a YouTube page that contained a video and user comments on the video (which are encoded in the <body> of the text as posts). The metadata contain a URL reference to the video and the YouTube channel that posted the video in <relatedItem> elements. The date when the page was created is not known. The example is adapted from the NottDeuYTSch corpus (?), where the video itself is not contained in the corpus.

```
<sourceDesc>
 <bibl>
  <title  type="main">Iron Man 3 in 3D (Official Trailer German) Parodie</title>
  <respStmt>
   <name  type="user">DieAussenseiter</name>
   <resp>posted video, created page</resp>
  </respStmt>
  <distributor>YouTube</distributor>
  <ptr  type="url"
```

```
     target="https://www.youtube.com/watch?v=T-WU_3-0UpU"/>
    <series>
     <title>DieAussenseiter's Channel</title>
     <ptr  target="https://www.youtube.com/watch?v=UCKn1vL4Ou4DKu0BlcK3NlDQ"/>
    </series>
  </bibl>
</sourceDesc>
```

The following example is the <sourceDesc> of a Wikipedia talk page. Note that a <relatedItem> element is used to record a reference to the Wikipedia article that the transcribed discussion is about.

```
<sourceDesc>
  <bibl>
    <title  type="main">Diskussion:FKM-Richtlinie</title>
    <author>
      <name  type="user">OnkelSchuppig</name>, et al.</author>
    <publisher>Wikimedia Foundation, Inc.</publisher>
    <ptr  target="https://de.wikipedia.org/wiki/Diskussion:FKM-Richtlinie"
      type="page_url"  targetLang="de"/>
    <date  type="last-change"
      when="2013-09-14T17:04:48Z"/>
    <idno  type="wikipedia-id">7632113</idno>
    <relatedItem  type="articleLink">
      <ref  n="5138958"
        target="https://de.wikipedia.org/wiki/FKM-Richtlinie"  targetLang="de">FKM-Richtlinie</ref>
    </relatedItem>
  </bibl>
</sourceDesc>
```

### 9.5.3 Documenting the Sampling of CMC data

The documentation of how the data were collected, e.g. how it was scraped or sampled from the web, or downloaded from a server, should be recorded in the <samplingDecl>. Like other metadata, information about sampling should be recorded at the highest level applicable. That is, if the information applies to an entire corpus, the <samplingDecl> should appear in the <teiHeader> of the corpus level; if the information is different for each document, it should appear in the <teiHeader> of the document level texts.

The sampling information typically considered of interest consists of at least the following four components:

- interface: The API that was used for the download, possibly encoded as a <name type="API">;
- client: The client or other tool that was used for the download, possibly encoded as a <name type="client">;
- query: The query or command used for the download, possibly encoded with a <ptr type="query"> when it is a URI, or a <code> when it is a command;
- date: The date of the download.

For example, in the case of an X (Twitter) corpus a sampling declaration might look like the following:

```
<samplingDecl>
  <p>Sampled using the <name  type="API">Twitter Filtered stream v2-API</name> (see
<ptr  type="APIdoc"
     target="https://developer.twitter.com/en/docs/twitter-api/tweets/filtered-stream/api-referen
Filtered for the German language and the following countries: Germany, Austria,
    Belgium, Switzerland, Denmark, and Luxembourg. Downloaded on <date  when="2022-12-12">Mon
12 Dec 22</date> using the command
  <code>requests.get("https://api.twitter.com/2/tweets/search/stream",
```

```
       headers=headers, params=params, stream=True,)</code> in the python script
<name type="script">collectFilteredTwitterStream.py</name>. </p>
</samplingDecl>
```

The <samplingDecl> of a Usenet Newsgroup corpus:

```
<samplingDecl>
  <p>Downloaded from the news.individual.de server on 2016-01-15 using nntp client in
     Python</p>
</samplingDecl>
```

### 9.5.4 Participants

A <listPerson> may be used to maintain an inventory of users and bots taking part in a CMC interaction, along with information about them. As with other such contextual information, it may be kept in the <teiHeader> (where it would occur in a <particDesc> within a <profileDesc>) or in a separate document completely. In either case, an encoded <post> may then be linked to its author by use of the *who* attribute.

In the following example, a list of participants is maintained in a <teiHeader>.

```
<!-- In the <teiHeader>: --><profileDesc>
  <particDesc>
    <listPerson>
      <person role="user" sex="male"
        xml:id="cmc_user_01">
        <persName type="userName">M</persName>
        <note type="link">/wiki/User:M</note>
        <affiliation>
          <email>mike@mydomain.com</email>
          <country>CH</country>
        </affiliation>
      </person>
<!-- … more persons … -->
      <person role="user" sex="female"
        xml:id="cmc_user_06">
        <persName type="userName">P</persName>
        <note type="link">/wiki/User:P</note>
        <affiliation>
          <email>pat@super.net</email>
          <country>ES</country>
        </affiliation>
      </person>
      <person role="user" xml:id="cmc_user_07">
        <persName type="userName">PKP</persName>
        <note type="link">/wiki/User:Pi</note>
      </person>
    </listPerson>
  </particDesc>
</profileDesc>
<!-- In the <body>: -->
<div type="wiki_discussion_page" n="073">
<!-- 4 other <post>s -->
  <post modality="written"
    xml:id="cmc_post04" indentLevel="1"
    replyTo="#cmc_post_073.004" who="#cmc_user_06">
```

```
    <p>Those haven't happened. If they do, we can revisit the concern.</p>
    <signed  generatedBy="template"
      rend="noLineBreak">
      <ref  target="/wiki/User:P">P</ref>
      <date>01:35, 8 April 2014 (UTC)</date>
    </signed>
  </post>
</div>
```

In the following version of the <body> portion of the same example, the list of interactants is stored in a separate file (in this case the file userList.xml in the same directory).

```
<!-- In the <body>: --><div  type="wiki_discussion_page"  n="073">
<!-- 4 other <post>s -->
  <post  modality="written"
    xml:id="cmc_post05"  indentLevel="1"
    replyTo="#cmc_post_073.004"  who="./userList.xml#cmc_user_06">
    <p>Those haven't happened. If they do, we can revisit the concern.</p>
    <signed  generatedBy="template"
      rend="noLineBreak">
      <ref  target="/wiki/User:P">P</ref>
      <date>01:35, 8 April 2014 (UTC)</date>
    </signed>
  </post>
</div>
```

Alternatively, a <prefixDef> may be used to declare a prefix which can be used in the value of *who* to generate a complete URI, thus making the values of *who* shorter, less error-prone, and easier to maintain. For example, the prefix `uL:` could be used to map the value uL:06 to `file:/userList.xml#cmc_user_06`. See *17.2.3. Using Abbreviated Pointers* for more information on establishing prefix definitions.

This indirection—using a <listPerson>, particularly one in a separate file, to store information about the users involved in a CMC interaction—is particularly useful when there is both a need to keep such information locally, and to remove it (e.g., to 'anonymize' the data) when the data are published or shared with other researchers.

### 9.5.5 Timeline

From most CMC environments, user posts come provided with a timestamp marking the time (often down to the second) when the post arrived and was registered at the CMC server. In the display of chat interactions, for instance, the time is automatically added by the system and usually precedes or follows the actual content of the post. In Wikipedia talk, a timestamp is automatically added when the user inserts his or her signature. A timestamp in the text body may be transcribed using a <date> or <time> element, in which case the *when* attribute may be used to record a normalized version of the date, time, or date and time if this information is available or reconstructible.

```
<post  modality="written"  rend="color:black"
  who="#f2213001.A06"  xml:id="cmc_post06">
  <time  generatedBy="system">21:52</time>
das ist auf jedenfall krankheit

</post>
```

```
<post modality="written"
  xml:id="cmc_post07" indentLevel="1" who="#u006" synch="#t006">
  <p>Those haven't happened. If they do, we can revisit the concern.</p>
  <signed generatedBy="template">
    <ref target="/wiki/User:P">P</ref>
    <date when="2014-04-08T01:35:00Z">01:35, 8 April 2014 (UTC)</date>
  </signed>
</post>
```

Alternatively the timestamp may be recorded using the *when* attribute of <post>. In this case, if the details of how the timestamp appeared in the original are considered unimportant, the actual transcription may be omitted.

```
<post modality="written"
  when="2014-04-08T01:35:00Z" who="#u006">
  <p>Those haven't happened. If they do, we can revisit the concern.</p>
  <signed generatedBy="template">
    <ref target="/wiki/User:P">P</ref>
  </signed>
</post>
```

Instead of transcribing timestamps or recording the timestamp directly on an attribute of <post>, all timestamps of a set of posts can be collected in <when> elements in a <timeline> element in the <teiHeader>, most suitably in the <interaction> element (itself in the <textDesc> in the <profileDesc>). In which case, similar to the encoding of transcripts of spoken utterances (for which see *8. Transcriptions of Speech*), each individual post can be linked to its timestamp via the *synch* attribute as in the following alternative encoding of the Wikipedia talk example above.

```
<profileDesc>
  <particDesc>
    <listPerson>
      <person role="user" xml:id="u001">
        <persName type="userName">M</persName>
        <note type="link">/wiki/User:M</note>
      </person>
<!-- more persons -->
      <person role="user" xml:id="u006">
        <persName type="userName">P</persName>
        <note type="link">/wiki/User:P</note>
      </person>
      <person role="user" xml:id="u007">
        <persName type="userName">PKP</persName>
        <note type="link">/wiki/User:Pi</note>
      </person>
    </listPerson>
  </particDesc>
  <textDesc>
    <channel/>
    <constitution/>
    <derivation/>
    <domain/>
    <factuality/>
    <interaction>
      <timeline>
        <when xml:id="t001"
          absolute="2011-03-23T19:56:00"/>
```

```
      <when  xml:id="t002"
        absolute="2011-06-14T21:22:00"/>
      <when  xml:id="t003"
        absolute="2011-06-14T23:28:00"/>
      <when  xml:id="t004"
        absolute="2011-07-02T07:20:00"/>
      <when  xml:id="t005"
        absolute="2011-07-21T01:00:00"/>
      <when  xml:id="t006"
        absolute="2014-04-08T01:35:00"/>
    </timeline>
  </interaction>
  <preparedness/>
  <purpose/>
 </textDesc>
</profileDesc>
```

Note that the *synch* attribute is provided by the module described in chapter *17. Linking, Segmentation, and Alignment*.

Removing timestamps from the text body can help meet requirements of text anonymization. For instance, if the <particDesc> and the <timeline> are stored in a separate file, the rest of the corpus can be distributed without this separate file. Thus the recipient of the corpus may know in what order posts were made (if the values of the *synch* are sequential), and will be able to group posts made by the same user, but will not have exact timestamps or actual user names, thus providing a significant degree of anonymization.

```
<post  modality="written"
 xml:id="cmc_post08"  indentLevel="1"  who="#u006"  synch="#t006">
 <p>Those haven't happened. If they do, we can revisit the concern. </p>
 <signed  generatedBy="template"> [_DELETED-SIGNATURE_]
 <date  synch="#t007">[_DELETED-TIMESTAMP_]</date>
 </signed>
</post>
```

As demonstrated above, the *synch* attribute can be used on <date> or <time> (or indeed any other element) rather than on the <post> itself.

## 9.6 Recommendations for Encoding CMC Microstructure

### 9.6.1 Emojis and Emoticons

Emojis are iconic or symbolic, invariant graphic units which the users of social media applications such as WhatsApp, Instagram, and X (Twitter) can select from a menu or 'emoji keyboard' and embed into their written posts. Examples are ❤, 😂, 😍, 🤔, and 🙏. An emoji is encoded by one or more Unicode characters which are intended to be mapped directly to a pictorial symbol.

Emoticons predate emojis and are created as combinations of ASCII punctuation and other characters using the keyboard. Examples are `:-)`, `;-)`, `:-(`, `:-x`, `\O/`, and `Oo`. They first occurred on a computer bulletin board system at Carnegie Mellon University (Fahlman, 2021) and then became frequent in chat communications during the mid-1980s. An emoticon typically consists of several Unicode characters (from the ASCII subset) in a row, each of which has an intended use other than as part of an emoticon.

Both emoticons and emojis may be simply transcribed as a sequence of characters. As with any other characters, they may be entered as numeric character entities if this is more convenient. (E.g., ❤ might be transcribed as `❤` in any XML document, including a TEI document; see *vi.2.4 Entry of Characters*.)

When the text of a post is being tokenized, e.g. for linguistic analysis, it may be useful to encode the emoticon or emoji as a separate token. In such cases elements such as <w> or <c> may be used for tokenization, and the *pos* attribute may be used to indicate that the encoded string is an emoji or an emoticon. (See *18.1. Linguistic Segment Categories*.)

For example, the source post ‘da bin ich nicht so empfindlich ;)’ (English:. ‘I am not so touchy with that ;)’) ends with an emoticon, and might be encoded as follows:

```
<post>
  <w  pos="ADV">da</w>
  <w  pos="VAFIN">bin</w>
  <w  pos="PPER">ich</w>
  <w  pos="PTKNEG">nicht</w>
  <w  pos="ADV">so</w>
  <w  pos="ADJD">empfindlich</w>
  <w  pos="EMOASC">;)</w>
</post>
```

Similarly, the source post ‘Klar ’ (‘Sure ’ in English) might be encoded as follows:

```
<post  modality="written"
  generatedBy="human"  synch="#lscLB.t004"  who="#lscLB.A03"
  xml:id="cmc_post21">
  <w  pos="ADV">Klar</w>
  <w  pos="EMOIMG"></w>
</post>
```

The values of *pos* in the above examples are from the STTS_IBK Tagset for German (see **?**), which includes tags for CMC-specific elements such as EMOASC for an ASCII-based emoticon and EMOIMG for an icon-based emoji.

Alternatively, e.g. when <w> is not regularly used to encode tokens in the TEI document, <c> may be used to mark an emoji. For example, the source post ‘Da kostet ein Haarschnitt 50 € ’ (from the corpus **?**, in English ‘A haircut there costs 50 € ’) might be encoded as follows:

```
<post  xml:lang="de">Da kostet ein Haarschnitt 50 € <c  type="emoji"  ana="#fsif"
    generatedBy="template"></c>
</post>
```

Sometimes, e.g. when the source of the TEI document was a web page in HTML, the emojis may occur only as an icon graphic in the source. In such a case, they may be encoded using <figure>. The corresponding Unicode character can then be recorded in the <desc> element by the encoder if desired.

For example, the source text: ‘… ich überlege noch ’ (English: ‘… I’m still thinking ’) might be encoded as follows:

```
<post  modality="written"
  generatedBy="human"  synch="#lscLB.t004"  who="#lscLB.A03"
  xml:id="cmc_post22"> ... ich überlege noch <figure  type="emoji"
    generatedBy="template">
    <graphic  url="fig1.png"/>
    <desc  type="gloss"  xml:lang="en">see no evil monkey</desc>
    <desc  type="unicode">U+1F648</desc>
  </figure>
</post>
```

### 9.6.2 Posts with Graphics

A post in a CMC interaction may contain a graphic in addition to some text or even contain only a graphic (without any text). As explained in *9.3.2. Attributes Specific to CMC post*, the modality of such a post should be considered as written. To encode the graphic information, the <figure> element may be used at the appropriate place.

In the following example a private chat post that contained only a screenshot of a google search result for a hairdresser is encoded as a <post> with a child <figure>. A link to the graphic file itself is not included presumably because this is a text-only corpus that does not include images.

```
<post modality="written"
  generatedBy="human" synch="#t005" who="#A06"
  xml:id="cmc_post23">
  <figure type="image" generatedBy="human">
    <desc>screenshot of the google search for hairdresser "Pasha's Haare'm" with the
        average google rating (4,5 of 5 stars), the address, the phone number, and the
        opening hours. </desc>
  </figure>
</post>
```

The following is an example of the encoding of a tweet which contains both text (including hashtags and mentions) and a graphic. The <graphic> element retains the URL of the graphic on the web just as in the source.

```
<post type="tweet" modality="written"
  generatedBy="human" synch="#tweetsbcrn18.t006"
  xml:id="cmc_post_1043823300479258624" who="#u1" xml:lang="de">
  <time generatedBy="system"> 16:24 </time> Bro Tri-Engel...so hab ich mir das
vorgestellt!!! @AndreLo79 #bcrn18 #wikidach @Heiko komm' mal Twitter! #Engel
<figure type="image" generatedBy="human">
    <graphic url="https://pbs.twimg.com/media/DnxnwN9XsAEHXw2.jpg:large"/>
  </figure>
</post>
```

### 9.6.3 Circulation

The following recommendations on how to encode features of the circulation of posts, such as IDs, re-posts (retweets), hashtags, and mentions use X (Twitter) posts (tweets) as an example; this phenomenon is not in any way unique to X (Twitter), however.

In the following example, the type of post (in this case, a tweet) is recorded using the *type* attribute of <post>. If it were useful to record a particular sub-categorization of tweet, the *subtype* attribute could also be used. Furthermore, the original unique identifer of the tweet as supplied by X (Twitter) is recorded as part of the value of the *xml:id* attribute of the <post>.

Also in the following example a retweet and its corresponding retweeted tweet are encoded as two separate posts each with its own set of attributes. The post representing the retweet itself does not contain or duplicate the content of the retweeted tweet. Instead it refers to the ID of the retweeted tweet via a <ptr> in the post content. All original content of the retweet goes in the content of the <post> element as well. In addition, the hashtags found in the body of the source tweets have been encoded using <ref> elements (with a *type* of hashtag), as they are links like any other hyperlink.

```
<post modality="written"
  generatedBy="human" type="tweet" who="#u1"
  xml:id="cmc_post_1043796550101716993" synch="#tweetsbcrn18.t004" xml:lang="de">
  <ptr type="retweet"
    target="#cmc_post_1043796093786566656"/> Ich mich auch? <ref type="hashtag"
    target="https://twitter.com/hashtag/dynamicduo?src=hash">#dynamicduo</ref>
  <ref type="hashtag"
    target="https://twitter.com/hashtag/wirk%C3%BCmmernunsauchumIhrenEmpfang?src=hash">#wirkümmern
  <ref type="hashtag"
    target="https://twitter.com/hashtag/bcrn18?src=hash">#bcrn18</ref>
  <ref type="hashtag"
```

```
   target="https://twitter.com/hashtag/wikidach?src=hash">#wikidach</ref>
</post>
<post modality="written"
  generatedBy="human" type="tweet" who="#u2"
  synch="#tweetsbcrn18.t003" xml:lang="de"
  xml:id="cmc_post_1043796093786566656">
  <time generatedBy="system"> 14:35 </time> Immer wieder gerne. Kann ich mich schon
für
nächstes Jahr als Empfangs- <ref type="hashtag"
   target="https://twitter.com/hashtag/Engel?src=hash">#Engel</ref> für das nächste
BarCamp bewerben <w pos="EMO">□</w>
  <ref type="hashtag"
   target="https://twitter.com/hashtag/bcrn18?src=hash">#bcrn18</ref>
  <trailer>
   <fs>
    <f name="favoritecount">
     <numeric value="4"/>
    </f>
   </fs>
  </trailer>
</post>
```

Note that in the above example 'CoMeRe' style (cf. ?) encoding is used to represent the number of favorites. It would also be reasonable to use a TEI <measure> element instead of the <fs>.

### 9.6.4 Linguistic Annotation

For encoding linguistic analyses of CMC text, we may use the dedicated elements and attributes from the analysis module, which is described in *17. Linking, Segmentation, and Alignment.* For example, the tokenization (segmentation into word-like units) of a CMC text should be encoded using the <w> element.

Let us take, for example a posting that contains the content '00:22 Bin soooooo im stress gewesen ich Armer lol' (in English: 'I was soooooo stressed out poor me lol'). This may be encoded as follows.

```
<post generatedBy="human"
  rend="color:black" synch="#t010.a" who="#A03.a"
  xml:id="cmc_post_m16.a">
  <time>00:22</time>
  <w>Bin</w>
  <w>soooooo</w>
  <w>im</w>
  <w>stress</w>
  <w>gewesen</w>
  <w>ich</w>
  <w>Armer</w>
  <w>lol</w>
</post>
```

In many CMC genres, especially in private chat, informal writing abounds including irregular spellings imitating spoken language, omitted word boundaries, and spurious boundaries leading to tokens separated in parts. For encoding these writing phenomena typical of CMC, the TEI attributes *norm* and *join* may be used.

For example, the normalized spelling of an irregularly spelled word may be recorded using the *norm* attribute (from `att.linguistic`):

```
<post generatedBy="human"
  rend="color:black" synch="#t010.b" who="#A03.b"
  xml:id="cmc_post_m16.b">
  <time> 00:22 </time>
  <w>Bin</w>
  <w norm="so">soooooo</w>
  <w>im</w>
  <w norm="Stress">stress</w>
  <w>gewesen</w>
  <w>ich</w>
  <w>Armer</w>
  <w>lol</w>
</post>
```

When the boundaries between <w> elements are generally thought of as denoting word boundaries, we can keep track of boundaries not present in the source by using the *join* attribute, also from `att.linguistic`. For example, for an original post that has nothing more than the token 'Inmyoffice', the following encoding demonstrates an interpretation that the single token represents the three words 'In my office':

```
<post>
  <w>In</w>
  <w join="left">my</w>
  <w join="left">office</w>
</post>
```

Alternatively, and especially when the normalization information pertains to more than one token, we can apply the notation using the elements <reg> and <orig>, related by a <choice> element as described in *3.5.2. Regularization and Normalization*.

```
<post>
  <choice>
    <orig>
      <w pos="VAPPER" lemma="">hastes</w>
    </orig>
    <reg>
      <w pos="VAFIN" lemma="haben">hast</w>
      <w pos="PPER" lemma="du">du</w>
      <w pos="PPER" lemma="es">es</w>
    </reg>
  </choice>
</post>
```

Other analysis attributes like *lemma* and *pos* (for part of speech) may be used as with traditional text. It is a matter of the tagset used to cater for POS categories that are appropriate for CMC. In the example below, for instance, the tag AKW stands for *Aktionswort* ( *action word*, see ?).

```
<post generatedBy="human"
  rend="color:black" synch="#t010.c" who="#A03.c"
  xml:id="cmc_post_m16.c">
  <time> 00:22 </time>
  <w lemma="sein" pos="VAFIN" xml:id="m16.t9">Bin</w>
  <w lemma="so" norm="so" pos="PTKIFG"
```

```
    xml:id="m16.t10">soooooo</w>
  <w lemma="in" pos="APPRART"
    xml:id="m16.t11">im</w>
  <w lemma="Stress" norm="Stress" pos="NN"
    xml:id="m16.t12">stress</w>
  <w lemma="sein" pos="VAPP" xml:id="m16.t13">gewesen</w>
  <w lemma="ich" pos="PPER" xml:id="m16.t14">ich</w>
  <w lemma="Armer" pos="NN" xml:id="m16.t15">Armer</w>
  <w lemma="lol" pos="AKW" xml:id="m16.t16">lol</w>
</post>
```

### 9.6.5 Named Entities and Anonymization

Named entities (NEs) may be marked up using <name> or the elements encoding different subcategories of names as described in *14. Names, Dates, People, and Places* such as <surname> or <geogName>, or <rs> for a general referencing string. In the following chat example (adapted from ?), nicknames are linked to a <person> entry as shown in section *9.5.4. Participants* via the *ref* attribute.

```
<post modality="written"
  generatedBy="system" rend="color:black" synch="#f2213001.t007"
  type="standard" who="#f2213001.A04"
  xml:id="f2213001.m27.eg35">
  <name ref="#f2213001.A04" type="NICK">
    <w lemma="Konstanze" pos="NE"
      xml:id="f2213001.m27.t1">Konstanze</w>
  </name>
  <w lemma="versuchen" pos="VVPP">versucht</w>
  <name ref="#f2213001.A03" type="NICK">
    <w lemma="Nasenloch" pos="NN">nasenloch</w>
  </name>
  <w lemma="die" pos="ART">den</w>
  <w lemma="Wunsch" pos="NN">wunsch</w>
  <w lemma="zu" pos="PTKZU">zu</w>
  <w lemma="erfüllen" pos="VVINF">erfüllen</w>
<!-- ... -->
</post>
```

In the following version of the same chat snippet, the text strings with the nicknames have been replaced by category label strings for the purpose of anonymization.

```
<post modality="written"
  generatedBy="system" rend="color:black"
  synch="#f2213001a.t007" type="standard" who="#f2213001a.A04"
  xml:id="f2213001a.m27.eg35">
  <name ref="#f2213001a.A04" type="NICK">
    <w pos="NE" xml:id="f2213001a.m27.t1">
      <gap reason="anonymization" unit="token"
        quantity="1"/>
      <supplied reason="anonymization">[_FEMALE-PARTICIPANT-A04_]</supplied>
    </w>
  </name>
  <w lemma="versuchen" pos="VVPP">versucht</w>
  <name ref="#f2213001a.A03" type="NICK">
    <w pos="NN">
```

```
      <gap reason="anonymization" unit="token"
        quantity="1"/>
      <supplied reason="anonymization">[_PARTICIPANT-A04_]</supplied>
    </w>
  </name>
  <w lemma="die" pos="ART">den</w>
  <w lemma="Wunsch" pos="NN">wunsch</w>
  <w lemma="zu" pos="PTKZU">zu</w>
  <w lemma="erfüllen" pos="VVINF">erfüllen</w>
<!-- ... -->
</post>
```

In the preceding example, pairs of a <gap> and a <supplied> element encode the fact that some substring has been removed and replaced with another string for anonymization purposes. Note that in this example, the <name> and the <w> elements and their attributes also provide some categorical information about what has been removed. Using <gap> and <supplied> to record the anonymization is especially recommendable when the original name or referencing string has been 'pseudonymized', i.e. replaced by a different referencing string of the same ontological category (such as replacing the female name *Konstanze* by the female name *Kornelia*.). In that case, the markup would be the only place where it can be seen that a pseudonymization has been carried out, as in the following version of the example.

```
<post modality="written"
  generatedBy="system" rend="color:black"
  synch="#f2213001p.t007" type="standard" who="#f2213001p.A04"
  xml:id="f2213001p.m27.eg35">
  <name ref="#f2213001p.A04" type="NICK">
    <w pos="NE" xml:id="f2213001p.m27.t1">
      <gap reason="pseudonymization"
        unit="token" quantity="1"/>
      <supplied reason="pseudonymization">Kornelia</supplied>
    </w>
  </name>
  <w lemma="versuchen" pos="VVPP">versucht</w>
<!-- the rest of the post -->
</post>
```

## 9.7 The TEI CMC Module

The module described in this chapter makes available the following components:

**Module cmc: Computer-mediated communication**

- *Elements defined*: post
- Macros defined: macro.specialPara.cmc

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema.*

# 10 Dictionaries

This chapter defines a module for encoding lexical resources of all kinds, in particular human-oriented monolingual and multilingual dictionaries, glossaries, and similar documents. The elements described here may also be useful in the encoding of computational lexica and similar resources intended for use by language-processing software; they may also be used to provide a rich encoding for wordlists, lexica, glossaries, etc. included within other documents. Dictionaries are most familiar in their printed form; however, increasing numbers of dictionaries exist also in electronic forms which are independent of any particular printed form, but from which various displays can be produced.

Both typographically and structurally, print dictionaries are extremely complex. Such lexical resources are moreover of interest to many communities with different and sometimes conflicting goals. As a result, many general problems of text encoding are particularly pronounced here, and more compromises and alternatives within the encoding scheme may be required in the future.[71] Two problems are particularly prominent.

First, because the structure of dictionary entries varies widely both among and within dictionaries, the simplest way for an encoding scheme to accommodate the entire range of structures actually encountered is to allow virtually any element to appear virtually anywhere in a dictionary entry. It is clear, however, that strong and consistent structural principles do govern the vast majority of conventional dictionaries, as well as many or most entries even in more 'exotic' dictionaries; encoding guidelines should include these structural principles. We therefore define two distinct elements for dictionary entries, one (<entry>) which captures the regularities of many conventional dictionary entries, and a second (<entryFree>) which uses the same elements, but allows them to combine much more freely. It is however recommended that <entry> be used in preference to <entryFree> wherever possible. These elements and their contents are described in sections *10.2. The Structure of Dictionary Entries*, *10.6. Unstructured Entries*, and *10.4. Headword and Pronunciation References*.

Second, since so much of the information in printed dictionaries is implicit or highly compressed, their encoding requires clear thought about whether it is to capture the precise typographic form of the source text or the underlying structure of the information it presents. Since both of these views of the dictionary may be of interest, it proves necessary to develop methods of recording both, and of recording the interrelationship between them as well. Users interested mainly in the printed format of the dictionary will require an encoding to be faithful to an original printed version. However, other users will be interested primarily in capturing the lexical information in a dictionary in a form suitable for further processing, which may demand the expansion or rearrangement of the information contained in the printed form. Further, some users wish to encode *both* of these views of the data, and retain the links between related elements of the two encodings. Problems of recording these two different views of dictionary data are discussed in section *10.5. Typographic and Lexical Information in Dictionary Data*, together with mechanisms for retaining both views when this is desired.

To deal with this complexity, and in particular to account for the wide variety of linguistic contexts within which a dictionary may be designed, it can be necessary to customize or change the schema by providing more restriction or possibly alternate content models for the elements defined in this chapter. Section *10.3.2. Grammatical Information* illustrates this with the provision of a closed set of values for grammatical descriptors.

This chapter contains a large number of examples taken from existing print dictionaries; in each case, the original source is identified. In presenting such examples, we have tried to retain the original typographic appearance of the example as well as presenting a suggested encoding for it. Where this has not been possible (for example in the display of pronunciation) we have adopted the transliteration found in the electronic edition of the *Oxford Advanced Learner's Dictionary*. Also, the middle dot in quoted entries is rendered with a full stop, while within the sample transcriptions hyphenation and syllabification points are indicated by a vertical bar |, regardless of their appearance in the source text.

## 10.1 Dictionary Body and Overall Structure

Overall, dictionaries have the same structure of front matter, body, and back matter familiar from other texts. In addition, this module defines <entry>, and <entryFree>, as component-level elements which can occur directly within a text division or the text body.

The following tags can therefore be used to mark the gross structure of a printed dictionary; the dictionary-specific tags are discussed further in the following section.

[71] We refer the reader to previous and current discussions of a common format for encoding lexical resources. For example, ?; ?; ?; ?; ?; ?; ?; and ?; ?.

`<text>` (text) contains a single text of any kind, whether unitary or composite, for example a poem or drama, a collection of essays, a novel, a dictionary, or a corpus sample.

`<front>` (front matter) contains any prefatory matter (headers, abstracts, title page, prefaces, dedications, etc.) found at the start of a document, before the main body.

`<body>` (text body) contains the whole body of a single unitary text, excluding any front or back matter.

`<back>` (back matter) contains any appendixes, etc. following the main part of a text.

`<div>` (text division) contains a subdivision of the front, body, or back of a text.

`<entry>` (entry) contains a single structured entry in any kind of lexical resource, such as a dictionary or lexicon.

`<entryFree>` (unstructured entry) contains a single unstructured entry in any kind of lexical resource, such as a dictionary or lexicon.

As members of the classes `att.entryLike` and `att.sortable`, <entry> and <entryFree> share the following attributes:

`att.entryLike` provides attributes used to distinguish different styles of dictionary entries.

`@type` indicates type of entry, in dictionaries with multiple types.

`att.sortable` provides attributes for elements in lists or groups that are sortable, but whose sorting key cannot be derived mechanically from the element content.

`@sortKey` supplies the sort key for this element in an index, list or group which contains it.

The front and back matter of a dictionary may well contain specialized material such as lists of common and proper nouns, grammatical tables, gazetteers, a 'guide to the use of the dictionary', etc. These should be tagged using elements defined elsewhere in these Guidelines, chiefly in the core module (chapter *3. Elements Available in All TEI Documents*) together with the specialized dictionary elements defined in this chapter.

The <body> element consists of a set of *entries*, optionally grouped into one or several  elements. These text divisions might, for example, correspond to sections for different letters of the alphabet, or to sections for different languages in a bilingual dictionary, as in the following example:

```
<body>
  <div>
    <head>English-French</head>
    <entry>
      <form>
        <orth>cat</orth>
      </form>
<!-- ... -->
    </entry>
    <entry>
      <form>
        <orth>dog</orth>
      </form>
<!-- ... -->
    </entry>
    <entry>
      <form>
        <orth>horse</orth>
      </form>
<!-- ... -->
    </entry>
  </div>
  <div>
    <head>French-English</head>
    <entry>
      <form>
        <orth>chat</orth>
```

```
      </form>
<!-- ... -->
    </entry>
    <entry>
      <form>
        <orth>chien</orth>
      </form>
<!-- ... -->
    </entry>
    <entry>
      <form>
        <orth>cheval</orth>
      </form>
<!-- ... -->
    </entry>
  </div>
</body>
```

In a print dictionary, the entries are typically typographically distinct entities, each headed by some morphological form of the lexical item described (the *headword*), and sorted in alphabetical order or (especially for non-alphabetic scripts) in some other conventional sequence. Dictionary entries should be encoded as distinct successive items, each marked as an <entry> or <entryFree> element. The *type* attribute may be used to distinguish different types of entries, for example main entries, related entries, run-on entries, or entries for cross-references, etc.

Some dictionaries provide distinct entries for homographs, on the basis of etymology, part-of-speech, or both, and typically provide a numeric superscript on the headword identifying the homograph number. In these cases each homograph should be encoded as a separate entry; an outer <entry> element, perhaps with a *type* attribute, may be used to group such successive homograph entries. In addition to a series of <entry> elements, the <entry> may contain a preliminary <form> group (see section *10.3.1. Information on Written and Spoken Forms*) when information about hyphenation, pronunciation, etc., is given only once for two or more homograph entries. If the homograph number is to be recorded, the global attribute *n* may be used for this purpose. In some dictionaries, homographs are treated in distinct parts of the same entry; in these cases, they may be separated by use of the <hom> element, for which see section *10.2.1. Hierarchical Levels*.

A sort key, given in the *sortKey* attribute, is often required for superentries and entries, especially in cases where the order of entries does not follow the local character-set collating sequence (as, for example, when an entry for '3D' appears at the place where 'three-D' would appear).

A dictionary with no internal divisions might thus have a structure like the following; an outer <entry> element with a *type* attribute value of homograph is shown grouping two homograph entries.

```
<body>
  <entry>
    <form>
      <orth>manifestation</orth>
<!-- demonstration -->
    </form>
  </entry>
  <entry>
    <form>
      <orth>émeute</orth>
<!-- riot -->
    </form>
  </entry>
  <entry type="homograph">
    <entry n="1">
      <form>
```

```
      <orth>grève</orth>
<!-- strike -->
     </form>
    </entry>
    <entry n="2">
     <form>
      <orth>grève</orth>
<!-- shore -->
     </form>
    </entry>
  </entry>
</body>
```

The following example demonstrates a possible encoding of a traditional root-based dictionary, which starts with the root as the main headword followed by full-fledged lexicographic entries of derived headwords.

```
<entry type="wordFamily" xml:lang="ar"
  xml:id="syj">
  <form type="root">
    <orth>سيج</orth>
  </form>
  <pc>:</pc>
<!-- To fence (verb) -->
  <entry type="mainEntry" xml:id="syj1">
    <form type="lemma">
      <orth>سَّيج</orth>
    </form>
    <sense xml:id="syj1_sense1">
      <cit type="example">
        <quote>الكرم</quote>
      </cit>
      <pc>:</pc>
      <def>لعج ه ل اجاسي</def>
    </sense>
    <pc>·</pc>
  </entry>
<!-- A fence (noun) -->
  <entry type="mainEntry" xml:id="syj2">
    <form type="lemma">
      <orth>السياج</orth>
    </form>
    <form type="inflected">
      <gramGrp>
        <gram type="number" value="plural">ج</gram>
      </gramGrp>
      <form type="variant">
        <orth>سَياجات</orth>
      </form>
      <lbl>و</lbl>
      <form type="variant">
        <orth>أسوجة</orth>
      </form>
      <lbl>و</lbl>
      <form type="variant">
        <orth>أسوجَة</orth>
      </form>
```

```
      <lbl>و</lbl>
      <form type="variant">
        <orth>سُوج</orth>
      </form>
    </form>
    <pc>:</pc>
    <sense xml:id="syj2_sense1">
      <def>طئاحلا</def>
    </sense>
    <pc>||</pc>
    <sense xml:id="syj2_sense2">
      <def>ام طيحأ ُهيب ىلع ءيش مرك الك و لخنلا</def>
    </sense>
  </entry>
  <pc>·</pc>
<!-- A kind of fish -->
  <entry type="mainEntry" xml:id="syj3">
    <form type="lemma">
      <orth>نا جيْسلا</orth>
    </form>
    <pc>(</pc>
    <usg type="domain" value="animal">ح</usg>
    <pc>)</pc>
    <pc>:</pc>
    <sense xml:id="syj3_sense1">
      <def>عون من كمسلا</def>
    </sense>
  </entry>
</entry>
```

## 10.2 The Structure of Dictionary Entries

A simple dictionary entry may contain information about the form of the word treated, its grammatical characterization, its definition, synonyms, or translation equivalents, its etymology, cross-references to other entries, usage information, and examples. These we refer to as the *constituent parts* or *constituents* of the entry; some dictionary constituents possess no internal structure, while others are most naturally viewed as groups of smaller elements, which may be marked in their own right. In some styles of markup, tags will be applied only to the low-level items, leaving the constituent groups which contain them untagged. We distinguish the class of *top-level constituents* of dictionary entries, which can occur directly within the <entry> element, from the class of *phrase-level* constituents, which can normally occur only within top-level constituents. The top-level constituents of dictionary entries are described in section *10.2.2. Groups and Constituents*, and documented more fully, together with their phrase-level sub-constituents, in section *10.3. Top-level Constituents of Entries*.

In addition, however, dictionary entries often have a complex hierarchical structure. For example, an entry may consist of two or more sub-parts, each corresponding to information for a different part-of-speech homograph of the headword. The entry (or part-of-speech homographs, if the entry is split this way) may also consist of senses, each of which may in turn be composed of two or more sub-senses, etc. Each sub-part, homograph entry, sense, or sub-sense we call a *level*; at any level in an entry, any or all of the constituent parts of dictionary entries may appear. The hierarchical levels of dictionary entries are documented in section *10.2.1. Hierarchical Levels*.

### 10.2.1 Hierarchical Levels

The outermost structural level of an entry is marked with the elements <entry> or <entryFree>. The <hom> element marks the subdivision of entries into homographs differing in their part-of-speech. The <sense> element marks the subdivision of entries and part-of-speech homographs into senses; this element nests recursively in order to provide for a hierarchy of sub-senses of any depth. It is recommended to use the <sense> element even for an entry that has only one sense to group together all parts of the definition relating to the word sense since

this leads to more consistent encoding across entries. All of these levels may each contain any of the constituent parts of an entry. A special case of hierarchical structure is represented by the <re> (related entry) element, which is discussed in section *10.3.6. Related Entries*. Finally, the element <dictScrap> may be used at any point in the hierarchy to delimit parts of the dictionary entry which are structurally anomalous, as further discussed in section *10.6. Unstructured Entries*.

`<entry>` (entry) contains a single structured entry in any kind of lexical resource, such as a dictionary or lexicon.

`<entryFree>` (unstructured entry) contains a single unstructured entry in any kind of lexical resource, such as a dictionary or lexicon.

`<hom>` (homograph) groups information relating to one homograph within an entry.

`<sense>` groups together all information relating to one word sense in a dictionary entry, for example definitions, examples, and translation equivalents.
`@level` gives the nesting depth of this sense.

`<dictScrap>` (dictionary scrap) encloses a part of a dictionary entry in which other phrase-level dictionary elements are freely combined.

For example, an entry with two senses will have the following structure:

```
<entry>
  <sense n="1"/>
  <sense n="2"/>
</entry>
```

An entry with two homographs, the first with two senses and the second with three (one of which has two subsenses), may have a structure like this:

```
<entry>
  <hom n="1">
    <sense n="1">
<!-- ... -->
    </sense>
    <sense n="2">
<!-- ... -->
    </sense>
  </hom>
  <hom n="2">
    <sense n="1">
      <sense n="a">
<!-- ... -->
      </sense>
      <sense n="b">
<!-- ... -->
      </sense>
    </sense>
    <sense n="2">
<!-- ... -->
    </sense>
    <sense n="3">
<!-- ... -->
    </sense>
  </hom>
</entry>
```

In some dictionaries, homographs have separate entries; in such a case, as noted in section *10.1. Dictionary Body and Overall Structure*, the two homographs may be treated as entries, optionally grouped in an outer <entry> element:

```
<entry>
  <entry  n="1"  type="hom">
    <sense  n="1">
<!-- ... -->
    </sense>
    <sense  n="2">
<!-- ... -->
    </sense>
  </entry>
  <entry  n="2"  type="hom">
    <sense  n="1">
      <sense  n="a">
<!-- ... -->
     </sense>
     <sense  n="b">
<!-- ... -->
     </sense>
    </sense>
    <sense  n="2">
<!-- ... -->
    </sense>
    <sense  n="3">
<!-- ... -->
    </sense>
  </entry>
</entry>
```

The hierarchic structure of a dictionary entry is enforced by the structures defined in this module. The content model for <entry> specifies that entries do not nest, that homographs nest within entries, and that senses nest within entries, homographs, or senses, and may be nested to any depth to reflect the embedding of sub-senses. Any of the top-level constituents (<def>, <usg>, <form>, etc.) can appear at any level (i.e., within entries, homographs, or senses).

### 10.2.2 Groups and Constituents

As noted above, dictionary entries, and subordinate levels within dictionary entries, may comprise several constituent parts, each providing a different type of information about the word treated. The *top-level constituents* of dictionary entries are:

- information about the form of the word treated (orthography, pronunciation, hyphenation, etc.)
- grammatical information (part of speech, grammatical sub-categorization, etc.)
- definitions or translations into another language
- etymology
- examples
- usage information
- cross-references to other entries
- notes
- entries (often of reduced form) for related words, typically called *related entries*

Any of the hierarchical levels (<entry>, <entryFree>, <hom>, and <sense>) may contain any of these top-level constituents, since information about word form, particular grammatical information, special pronunciation, usage information, etc., may apply to an entire entry, or to only one homograph, or only to a particular sense. The examples below illustrate this point.

The following elements are used to encode these top-level constituents:

`<form>` (form information group) groups all the information on the written and spoken forms of one headword.

`<gramGrp>` (grammatical information group) groups morpho-syntactic information about a lexical item, e.g. <pos>, <gen>, <number>, <case>, or <iType> (inflectional class).

`<def>` (definition) contains definition text in a dictionary entry.

`<cit>` (cited quotation) contains a quotation from some other document, together with a bibliographic reference to its source. In a dictionary it may contain an example text with at least one occurrence of the word form, used in the sense being described, or a translation of the headword, or an example.

`<usg>` (usage) contains usage information in a dictionary entry.

`<xr>` (cross-reference phrase) contains a phrase, sentence, or icon referring the reader to some other location in this or another text.

`<etym>` (etymology) encloses the etymological information in a dictionary entry.

`<re>` (related entry) contains a dictionary entry for a lexical item related to the headword, such as a compound phrase or derived form, embedded inside a larger entry.

`<note>` (note) contains a note or annotation.

In a simple entry with no internal hierarchy, all top-level constituents can appear as children of <entry>.

> **com.peti.tor** `/k@m"petit@(r)/` n person who competes. OALD

```
<entry>
  <form>
    <orth>competitor</orth>
    <hyph>com|peti|tor</hyph>
    <pron>k@m"petit@(r)</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
  <def>person who competes.</def>
</entry>
```

For the elements which appear within the <form> and <gramGrp> elements of this and other examples, see below, section *10.3.1. Information on Written and Spoken Forms*, and section *10.3.2. Grammatical Information*.

Any top-level constituent can appear at any level when the hierarchical structure of the entry is more complex. The most obvious examples are <def> and <cit>, which appear at the <sense> level when several senses or translations exist:

> **disproof** `(dɪs'pru:f)` *n* **1** facts that disprove something **2** the act of disproving. CED

```
<entry>
  <form>
    <orth>disproof</orth>
    <pron  notation="ipa">dɪs'pru:f</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
  <sense  n="1">
    <def>facts that disprove something</def>
  </sense>
  <sense  n="2">
    <def>the act of disproving</def>
  </sense>
</entry>
```

For ease of processing of such entries containing multiple senses along with those containing only a single sense, it is recommended to use <sense> in all entries to wrap those elements relating to a particular word sense.

In the following example, <gramGrp> is used to distinguish two homographs:

> **bray** /breI/ n cry of an ass; sound of a trumpet. · vt [VP2A] make a cry or sound of this kind. OALD

```
<entry>
  <form>
    <orth>bray</orth>
    <pron>breI</pron>
  </form>
  <hom>
    <gramGrp>
      <pos>n</pos>
    </gramGrp>
    <sense>
      <def>cry of an ass; sound of a trumpet.</def>
    </sense>
  </hom>
  <hom>
    <gramGrp>
      <pos>vt</pos>
      <subc>VP2A</subc>
    </gramGrp>
    <sense>
      <def>make a cry or sound of this kind.</def>
    </sense>
  </hom>
</entry>
```

Information of the same kind can appear at different levels within the same entry; here, grammatical information occurs both at entry and homograph level.

> **ca.reen** /k@"ri:n/ vt,vi 1 [VP6A] turn (a ship) on one side for cleaning, repairing, etc. 2 [VP6A, 2A] (cause to) tilt, lean over to one side. OALD

```
<entry>
  <form>
    <orth>careen</orth>
    <hyph>ca|reen</hyph>
    <pron>k@"ri:n</pron>
  </form>
  <gramGrp>
    <pos>vt</pos>
    <pos>vi</pos>
  </gramGrp>
  <sense  n="1">
    <gramGrp>
      <subc>VP6A</subc>
    </gramGrp>
    <def>turn (a ship) on one side for cleaning, repairing, etc.</def>
  </sense>
  <sense  n="2">
    <gramGrp>
      <subc>VP6A</subc>
```

```
      <subc>VP2A</subc>
    </gramGrp>
    <def>(cause to) tilt, lean over to one side.</def>
  </sense>
</entry>
```

Alone among the constituent groups, <form> can appear at the <entry>, <hom>, and <sense> levels:

> **a.ban.don** 1/@"band@n/ v [T1] 1 to leave completely and for ever; desert: The sailors abandoned the burning ship. 2 …**abandon** 2 n [U] the state when one's feelings and actions are uncontrolled; freedom from control…LDOCE

```
<entry>
  <form>
    <orth>abandon</orth>
    <hyph>a|ban|don</hyph>
    <pron>@"band@n</pron>
  </form>
  <entry  n="1">
    <gramGrp>
      <pos>v</pos>
      <subc>T1</subc>
    </gramGrp>
    <sense  n="1">
      <def>to leave completely and for ever … </def>
    </sense>
    <sense  n="2"/>
  </entry>
  <entry  n="2">
    <gramGrp>
      <pos>n</pos>
      <subc>U</subc>
    </gramGrp>
    <sense>
      <def>the state when one's feelings and actions are uncontrolled; freedom
            from control…</def>
    </sense>
  </entry>
</entry>
```

## 10.3 Top-level Constituents of Entries

This section describes the top-level constituents of dictionary entries, together with the phrase-level constituents peculiar to each.

- the <form> element, which groups orthographic information and pronunciations, is described in section *10.3.1. Information on Written and Spoken Forms*
- the <gramGrp> element, which groups elements for the grammatical characterization of the headword, is described in section *10.3.2. Grammatical Information*
- the <def> element, which describes the meaning of the headword, is described in section *10.3.3. Sense Information*
- the <etym> element and its special phrase-level elements are documented in section *10.3.4. Etymological Information*
- the <cit> element and its specific applications are described in section *10.3.3. Sense Information* and section *10.3.5. Other Information*

- the <usg>, <lbl>, <xr>, and <note> elements are described in section *10.3.5. Other Information*
- the <re> element, which marks nested entries for related words, is described in section *10.3.6. Related Entries*

### 10.3.1 Information on Written and Spoken Forms

Dictionary entries most often begin with information about the form of the word to which the entry applies. Typically, the orthographic form of the word, sometimes marked for syllabification or hyphenation, is the first item in an entry. Other information about the word, including variant or alternate forms, inflected forms, pronunciation, etc., is also often given.

The following elements should be used to encode this information: the <form> element groups one or more occurrences of any of them; it can also be recursively nested to reflect more complex sub-grouping of information about word form(s), as shown in the examples.

`<form>` (form information group) groups all the information on the written and spoken forms of one headword.
- `@type` classifies form as simple, compound, etc.

`<orth>` (orthographic form) gives the orthographic form of a dictionary headword.
- `@type` gives the type of spelling.
- `@extent [att.partials]` indicates whether the pronunciation or orthography applies to all or part of a word.

`<pron>` (pronunciation) contains the pronunciation(s) of the word.
- `@extent [att.partials]` indicates whether the pronunciation or orthography applies to all or part of a word.

`<hyph>` (hyphenation) contains a hyphenated form of a dictionary headword, or hyphenation information in some other form.

`<syll>` (syllabification) contains the syllabification of the headword.

`<stress>` (stress) contains the stress pattern for a dictionary headword, if given separately.

`<lbl>` (label) contains a label for a form, example, translation, or other piece of information, e.g. abbreviation for, contraction of, literally, approximately, synonyms:, etc.

In addition to those listed above, the following elements, which encode morphological details of the form, may also occur within <form> elements:

`<gram>` (grammatical information) within an entry in a dictionary or a terminological data file, contains grammatical information relating to a term, word, or form.
- `@type` classifies grammatical information provided by <gram> according to some convenient and shared typology, ideally one defined in an external reference taxonomy, such as the CLARIN Concept Registry.

`<gen>` (gender) identifies the morphological gender of a lexical item, as given in the dictionary.

`<number>` (number) indicates grammatical number associated with a form, as given in a dictionary.

`<case>` (case) contains grammatical case information given by a dictionary for a given form.

`<per>` (person) contains an indication of the grammatical person (1st, 2nd, 3rd, etc.) associated with a given inflected form in a dictionary.

`<tns>` (tense) indicates the grammatical tense associated with a given inflected form in a dictionary.

`<mood>` (mood) contains information about the grammatical mood of verbs (e.g. indicative, subjunctive, imperative).

`<iType>` (inflectional class) indicates the inflectional class associated with a lexical item.
- `@type` indicates the type of indicator used to specify the inflection class, when it is necessary to distinguish between the usual abbreviated indications (e.g. *inv*) and other kinds of indicators, such as special codes referring to conjugation patterns, etc.

`<pos>` (part of speech) indicates the part of speech assigned to a dictionary headword such as noun, verb, or adjective.

`<subc>` (subcategorization) contains subcategorization information (transitive/intransitive, countable/non-countable, etc.)

`<colloc>` (collocate) contains any sequence of words that co-occur with the headword with significant frequency.

Of these, the <gram> element is most general, and all of the others are synonymous with a <gram> element with appropriate values (gen, number, case, etc.) for the *type* attribute.

The use of these elements as children of <form> is deprecated; instead, they should always be children of a <gramGrp> within <form> when describing that particular form of the word.

Different dictionaries use different means to mark hyphenation, syllabification, and stress, and they often use some unusual glyphs (e.g., the 'middle dot' for hyphenation). All of these glyphs are in the Unicode character set, as discussed in *v.7.1 Character References*. When transcribing representations of pronunciation the International Phonetic Alphabet should be used. It may be convenient (as has been done in the text of this chapter) to use a simple transliteration scheme for this; such a scheme should however be properly documented in the header.

In the simplest case, nothing is given but the orthography:

```
<form>
  <orth>doom-laden</orth>
</form>
```

Often, however, pronunciation is given.

**soucoupe** [sukup] ... DNT

```
<form>
  <orth>soucoupe</orth>
  <pron>sukup</pron>
</form>
```

For a variety of reasons including ease of processing, it may be desired to split into separate elements information which is collapsed into a single element in the source text; orthography and hyphenation may for example be transcribed as separate elements, although given together in the source text. For a discussion of the issues involved, and of methods for retaining both the presentation form and the interpreted form, see section *10.5. Typographic and Lexical Information in Dictionary Data*.

This example splits orthography and hyphenation, and adds syllabification because it differs from hyphenation:

**ar.ea** ... W7

```
<form>
  <orth>area</orth>
  <hyph>ar|ea</hyph>
  <syll>ar|e|a</syll>
</form>
```

Multiple orthographic forms may be given, e.g. to illustrate a word's inflectional pattern:

**brag** ... *vb* **brags**, **bragging**, **bragged** ... CED

```
<form>
  <orth>brag</orth>
</form>
<gramGrp>
  <pos>vb</pos>
</gramGrp>
```

```
<form  type="inflected">
  <orth>brags</orth>
  <orth>bragging</orth>
  <orth>bragged</orth>
</form>
```

Or the inflectional pattern may be indicated by reference to a table of paradigms, as here:

**horrifier** `[ORifje]` (7) vt ... [C/R]

```
<form>
  <orth>horrifier</orth>
  <pron>ORifje</pron>
  <gramGrp>
    <iType  type="vbtable">7</iType>
<!-- ... -->
  </gramGrp>
</form>
<!-- ... -->
```

Explanatory labels may be attached to alternate forms:

**MTBF** *abbreviation for* mean time between failures CED

```
<entry>
  <form  type="abbrev">
    <orth>MTBF</orth>
  </form>
  <form  type="full">
    <lbl>abbreviation for</lbl>
    <orth>mean time between failures</orth>
  </form>
</entry>
```

When multiple orthographic forms are given, a pronunciation may be associated with all of them, as here:

**biryani** *or* **biriani** `(ˌbɪrɪ'a:nɪ)` ... CED

```
<form>
  <orth>biryani</orth>
  <orth>biriani</orth>
  <pron  notation="ipa">ˌbɪrɪ'a:nɪ</pron>
</form>
```

In other cases, different pronunciations are provided for different orthographic forms; here, the <form> element is repeated to associate the first orthographic form explicitly with the first pronunciation, and the second orthographic form with the second pronunciation:

**mackle** `('mækᵊl)` *or* **macule** `('mækju:l)` ... CED

```
<form>
  <orth>mackle</orth>
  <pron  notation="ipa">'mækᵊl</pron>
</form>
<form>
  <orth>macule</orth>
  <pron  notation="ipa">'mækju:l</pron>
</form>
```

Recursive nesting of the <form> element can preserve relations among elements that are implicit in the text. For example, in the CED entry for 'hospitaller', it is clear that 'U.S.' is associated only with 'hospitaler', but that the pronunciation applies to both forms. The following encoding preserves these relations:

**hospitaller** *or US* **hospitaler** ('hɒspɪtələ) ... CED

```
<form>
  <orth>hospitaller</orth>
  <form>
    <usg  type="geo">US</usg>
    <orth>hospitaler</orth>
  </form>
  <pron  notation="ipa">'hɒspɪtələ</pron>
</form>
```

### 10.3.2 Grammatical Information

The <gramGrp> element groups grammatical information, such as part of speech, subcategorization information (e.g., syntactic patterns for verbs, count/mass distinctions for nouns), etc. It can contain any of the morphological elements defined in section *10.3.1. Information on Written and Spoken Forms* for <form> and can appear as a child of <entry>, <form>, <sense>, <cit>, or any other element containing content about which there is grammatical information. For example, in the entry '**pinna** ('pɪnə) *n*, *pl* **-nae** (-ni:) *or* **-nas** CED', the word defined can be either singular or plural; the 'pl.' specification applies only to the inflected forms provided. Compare this with 'pants (paents) pl. n.', where 'pl.' applies to the headword itself.

As noted above in section *10.3.1. Information on Written and Spoken Forms*, the elements for morphological information are simply shorthand for the general purpose <gram> element. Consider this entry for the French word *médire*:

**médire** v.t. ind. (de) ... PLC

This entry can be tagged using specialized grammatical elements:

```
<form>
  <orth>médire</orth>
</form>
<gramGrp>
  <pos>v</pos>
  <subc>t ind</subc>
  <colloc>de</colloc>
</gramGrp>
```

Or using the <gram> element:

```
<form>
  <orth>médire</orth>
</form>
<gramGrp>
  <gram  type="pos">v</gram>
  <gram  type="subc">t ind</gram>
  <gram  type="collocPrep">de</gram>
</gramGrp>
```

Like <form>, <gramGrp> can be repeated, recursively nested, or used at the <sense> level to show relations among elements.

**isotope** adj. et n. m. ... DNT

```
<form>
  <orth>isotope</orth>
</form>
<gramGrp>
  <pos>adj</pos>
</gramGrp>
<gramGrp>
  <pos>n</pos>
  <gen>m</gen>
</gramGrp>
```

**wits** (`wɪts`) *pl n* **1** (*sometimes singular*) the ability to reason and act, esp quickly ... CED

```
<entry>
  <form>
    <orth>wits</orth>
    <pron  notation="ipa">wɪts</pron>
  </form>
  <gramGrp>
    <number>pl</number>
    <pos>n</pos>
  </gramGrp>
  <sense  n="1">
    <gramGrp>
      <number>sometimes singular</number>
    </gramGrp>
    <def>the ability to reason and act, esp quickly …</def>
  </sense>
</entry>
```

### 10.3.3 Sense Information

Dictionaries may describe the meanings of words in a wide variety of different ways—by means of synonyms, paraphrases, translations into other languages, formal definitions in various highly stylized forms, etc. No attempt is made here to distinguish all the different forms which sense information may take; all of them may be tagged using the <def> element described in section *10.3.3.1. Definitions*.

As a special case it is frequently desirable to distinguish the provision of translation equivalents in other languages from other forms of sense information; the use of <cit type="translation"> (which groups a translation equivalent with related information such as its grammatical description) for this purpose is described in section *10.3.3.2. Translation Equivalents*.

#### 10.3.3.1 Definitions

Dictionary definitions are those pieces of prose in a dictionary entry that describe the meaning of some lexical item. Most often, definitions describe the headword of the entry; in some cases, they describe translated texts, examples, etc.; see <cit type="translation">, section *10.3.3.2. Translation Equivalents*, and <cit type="example">, section *10.3.5.1. Examples*. The <def> element directly contains the text of the definition; unlike <form> and <gramGrp>, it does not serve solely to group a set of smaller elements. The close analysis of definition text, such as the tagging of hypernyms, typical objects, etc., is not covered by these Guidelines.

Definitions may occur directly within an entry; when multiple definitions are given, they are typically identified as belonging to distinct senses, as here:

> **demigod** (…) n. 1.a. a being who is part mortal, part god. b. a lesser deity. 2. a godlike person. CP

```
<entry>
  <form>
    <orth>demigod</orth>
    <pron> … </pron>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
  <sense  n="1">
    <sense  n="a">
      <def>a being who is part mortal, part god.</def>
    </sense>
    <sense  n="b">
      <def>a lesser deity.</def>
    </sense>
  </sense>
  <sense  n="2">
    <def>a godlike person.</def>
  </sense>
</entry>
```

In multilingual dictionaries, it is sometimes possible to distinguish translation equivalents from definitions proper; here a <def> element is distinguished from the translation information within which it appears.

> **rémoulade** [Remulad] nf remoulade, rémoulade (*dressing containing mustard and herbs*). CR

```
<entry>
  <form>
    <orth>rémoulade</orth>
    <pron>Remulad</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
    <gen>f</gen>
  </gramGrp>
  <cit  type="translation"  xml:lang="en">
    <quote>remoulade</quote>
    <quote>rémoulade</quote>
    <def>dressing containing mustard and herbs</def>
  </cit>
</entry>
```

#### 10.3.3.2 Translation Equivalents

Multilingual dictionaries contain information about translations of a given word in some source language for one or more target languages. Minimally, the dictionary provides the corresponding translation in the target language; other material, such as morphological information (gender, case), various kinds of usage restrictions, etc., may also be given. If translation equivalents are to be distinguished from other kinds of sense information, they may be encoded using <cit type="translation">. The global *xml:lang* attribute should be used to specify the target language.

As in monolingual dictionaries, the <sense> element is used in multilingual dictionaries to group information (forms, grammatical information, usage, translation(s), etc.) about a given sense of a word where necessary. Information about the individual translation equivalents within a sense is grouped using <cit type="translation">. This information may include the translation text (tagged <q> or <quote>), morphological information (<gen>, <case>, etc.), usage notes (<usg>), translation labels (<lbl>), and definitions (<def>).When bibliographic data is provided, the <quote> element should be used.

`<cit>` (cited quotation) contains a quotation from some other document, together with a bibliographic reference to its source. In a dictionary it may contain an example text with at least one occurrence of the word form, used in the sense being described, or a translation of the headword, or an example.

`<lbl>` (label) contains a label for a form, example, translation, or other piece of information, e.g. abbreviation for, contraction of, literally, approximately, synonyms:, etc.

Note how in the following example, different translation equivalents are grouped into the same or different senses, following the punctuation of the source and the usage labels:

> **dresser** ... (a) (Theat) habilleur m, -euse f; (Comm: window ~) étalagiste mf. she's a stylish ~elle s'habille avec chic; V hair. (b) (tool) (for wood) raboteuse f; (for stone) rabotin m. CR

```
<entry n="1">
  <form>
    <orth>dresser</orth>
  </form>
  <sense n="a">
    <sense>
      <usg type="dom">Theat</usg>
      <cit type="translation" xml:lang="fr">
        <quote>habilleur</quote>
        <gramGrp>
          <gen>m</gen>
        </gramGrp>
      </cit>
      <cit type="translation" xml:lang="fr">
        <quote>-euse</quote>
        <gramGrp>
          <gen>f</gen>
        </gramGrp>
      </cit>
    </sense>
    <sense>
      <usg type="dom">Comm</usg>
      <form type="compound">
        <orth>window <oRef/>
        </orth>
      </form>
      <cit type="translation" xml:lang="fr">
        <quote>étalagiste</quote>
        <gramGrp>
          <gen>mf</gen>
        </gramGrp>
      </cit>
```

```
    </sense>
    <cit type="example">
      <quote>she's a stylish <oRef/>
      </quote>
      <cit type="translation" xml:lang="fr">
        <quote>elle s'habille avec chic</quote>
      </cit>
    </cit>
    <xr type="see">V. <ref target="#hair">hair</ref>
    </xr>
  </sense>
  <sense n="b">
    <usg type="category">tool</usg>
    <sense>
      <usg type="hint">for wood</usg>
      <cit type="translation" xml:lang="fr">
        <quote>raboteuse</quote>
        <gramGrp>
          <gen>f</gen>
        </gramGrp>
      </cit>
    </sense>
    <sense>
      <usg type="hint">for stone</usg>
      <cit type="translation" xml:lang="fr">
        <quote>rabotin</quote>
        <gramGrp>
          <gen>m</gen>
        </gramGrp>
      </cit>
    </sense>
  </sense>
</entry>
<!-- ... -->
<entry xml:id="hair">
  <sense>
<!-- ... -->
  </sense>
</entry>
```

In the following example, a distinction is made between the translation equivalent ('OAS') and a descriptive phrase providing further information for the user of the dictionary.

> **O.A.S.** ... nf (abrév de **Organisation de l'Armée secrète**) OAS (*illegal military organization supporting French rule of Algeria*). CR

```
<entry>
<!-- ... -->
  <cit type="translation" xml:lang="en">
    <quote>OAS</quote>
    <def>illegal military organization supporting French rule of
        Algeria</def>
  </cit>
</entry>
```

Note that <cit type="translation"> may also be used in monolingual dictionaries when a translation is given for a foreign word:

***havdalah*** *or* ***havdoloh*** *Hebrew* (havda'la; *Yiddish* hav'dɔlə) *n Judaism* the ceremony marking the end of the sabbath or of a festival, including the blessings over wine, candles and spices [literally: separation] CED

```
<entry type="foreign">
  <form>
    <orth>havdalah</orth>
    <orth>havdoloh</orth>
    <gramGrp>
      <gram type="pos">n</gram>
    </gramGrp>
  </form>
  <sense>
    <usg type="dom">Judaism</usg>
    <def>the ceremony marking the end of the sabbath or of a festival,
        including the blessings over wine, candles and spices</def>
  </sense>
  <cit type="translation" xml:lang="en">
    <usg type="style">literally</usg>
    <quote>separation</quote>
  </cit>
</entry>
```

### 10.3.4 Etymological Information

The element <etym> marks a block of etymological information. Etymologies may contain highly structured lists of words in an order indicating their descent from each other, but often also include related words and forms outside the direct line of descent, for comparison. Not infrequently, etymologies include commentary of various sorts, and can grow into short (or long!) essays with prose-like structure. This variation in structure makes it impracticable to define tags which capture the entire intellectual structure of the etymology or record the precise interrelation of all the words mentioned. It is, however, feasible to mark some of the more obvious phrase-level elements frequently found in etymologies, using tags defined in the core module or elsewhere in this chapter. Of particular relevance for the markup of etymologies are:

`<etym>` (etymology) encloses the etymological information in a dictionary entry.

`<lang>` (language name) contains the name of a language mentioned in etymological or other linguistic discussion.

`<date>` (date) contains a date in any format.

`<mentioned>` marks words or phrases mentioned, not used.

`<gloss>` (gloss) identifies a phrase or word used to provide a gloss or definition for some other word or phrase.

`<pron>` (pronunciation) contains the pronunciation(s) of the word.

`<usg>` (usage) contains usage information in a dictionary entry.

`<lbl>` (label) contains a label for a form, example, translation, or other piece of information, e.g. abbreviation for, contraction of, literally, approximately, synonyms:, etc.

As in other prose, individual word forms mentioned in an etymological description are tagged with <mentioned> elements. Pronunciations, usage labels, and glosses can be tagged using the <pron>, <usg>, and <gloss> elements defined elsewhere in these Guidelines. In addition, the <lang> element may be used to identify a particular language name where it appears, in addition to using the *xml:lang* attribute of the <mentioned> element.

Examples:

**abismo** m. (del gr. a priv. y byssos, fondo). Sima, gran profundidad. ...

```
<entry>
  <form>
    <orth>abismo</orth>
  </form>
  <etym>del <lang>gr.</lang>
    <mentioned>a</mentioned> priv. y <mentioned>byssos</mentioned>,
  <gloss>fondo</gloss>
  </etym>
<!-- ... -->
</entry>
```

**neume** \'n(y)üm\ n [F, fr. ML pneuma, neuma, fr. Gk pneuma breath — more at **pneumatic**]: any of various symbols used in the notation of Gregorian chant ... [WNC]

```
<entry>
<!-- ... -->
  <etym>
    <lang>F</lang> fr. <lang>ML</lang>
    <mentioned>pneuma</mentioned>
    <mentioned>neuma</mentioned> fr. <lang>Gk</lang>
    <mentioned>pneuma</mentioned>
    <gloss>breath</gloss>
    <xr  type="etym">more at <ptr  target="#pneumatic"/>
    </xr>
  </etym>
  <sense>
    <def>any of various symbols used in the notation of Gregorian chant
<!-- ... -->
    </def>
  </sense>
</entry>
<!-- ... -->
<entry  xml:id="pneumatic">
  <etym>
<!-- ... -->
  </etym>
</entry>
```

### 10.3.5 Other Information

#### 10.3.5.1 Examples

Dictionaries typically include examples of word use, usually accompanying definitions or translations. In some cases, the examples are quotations from another source, and are occasionally followed by a citation to the author.

The <cit type="example"> element contains usage examples and associated information; the example text itself should be enclosed in a <q> or <quote> element. The <cit> element associates a quotation with a bibliographic reference to its source.

`<q>` (quoted) contains material which is distinguished from the surrounding text using quotation marks or a similar method, for any one of a variety of reasons including, but not limited to: direct speech or thought, technical terms or jargon, authorial distance, quotations from elsewhere, and passages that are mentioned but not used.

`<quote>` (quotation) contains a phrase or passage attributed by the narrator or author to some agency external to the text.

`<cit>` (cited quotation) contains a quotation from some other document, together with a bibliographic reference to its source. In a dictionary it may contain an example text with at least one occurrence of the word form, used in the sense being described, or a translation of the headword, or an example.

Examples frequently abbreviate the headword, and so their transcription will frequently make use of the <oRef> element described below in section *10.4. Headword and Pronunciation References*.

Examples:

> **multiplex /…/** adj tech having many parts: the multiplex eye of the fly. LDOCE

```
<quote>the multiplex eye of the fly.</quote>
```

Or when one wants a more comprehensive representation of examples:

```
<cit  type="example">
  <quote>the multiplex eye of the fly.</quote>
</cit>
```

As the following example shows, <cit> can also contain elements such as <pron>, <def>, etc.

> **some** … 4. (*S*~ and *any* are used with *more*): Give me ~more/s@'m0:(r)/ OALD

```
<sense  n="4">
  <usg  type="colloc">
    <oRef  type="cap"/> and <mentioned>any</mentioned> are used with
  <mentioned>more</mentioned>
  </usg>
  <cit  type="example">
    <quote>Give me <oRef/> more</quote>
    <pron  extent="part">s@'m0:(r)</pron>
  </cit>
</sense>
```

In multilingual dictionaries, examples may also be accompanied by translations:

> **horrifier** … vt to horrify. **elle était horrifiée par la dépense** she was horrified at the expense. CR

```
<entry>
<!-- ... -->
  <cit  type="translation"  xml:lang="en">
    <quote>to horrify</quote>
  </cit>
  <cit  type="example">
    <quote>elle était horrifiée par la dépense</quote>
    <cit  type="translation"  xml:lang="en">
      <quote>she was horrified at the expense.</quote>
    </cit>
  </cit>
</entry>
```

When a source is indicated, the example should be marked with a <bibl> element:

> **valeur** … n. f. … 2. Vx. Vaillance, bravoure (spécial., au combat). ‘La valeur n’attend pas le nombre des années’ (Corneille). … DNT

```
<sense n="2">
  <usg type="time">Vx.</usg>
  <def>Vaillance, bravoure (spécial., au combat)</def>
  <cit type="example">
    <quote>La valeur n'attend pas le nombre des années</quote>
    <bibl>
      <author>Corneille</author>
    </bibl>
  </cit>
</sense>
```

#### 10.3.5.2 Usage Information and Other Labels

Most dictionaries provide restrictive labels and phrases indicating the usage of given words or particular senses. Other phrases, not necessarily related to usage, may also be attached to forms, translations, cross-references, and examples. The following elements are provided to mark up such labels:

`<usg>` (usage) contains usage information in a dictionary entry.

`<lbl>` (label) contains a label for a form, example, translation, or other piece of information, e.g. abbreviation for, contraction of, literally, approximately, synonyms:, etc.

As indicated in the following section (*10.3.5.3. Cross-References to Other Entries*), the <lbl> element may be used for any kind of significative phrase or label within the text. The <usg> element is a specialization of this to mark usage labels in particular. Usage labels typically indicate

- temporal use (archaic, obsolete, etc.)
- register (slang, formal, taboo, ironic, facetious, etc.)
- style (literal, figurative, etc.)
- connotative effect (e.g. derogatory, offensive)
- subject field (Astronomy, Philosophy, etc.)
- national or regional use (Australian, U.S., Midland dialect, etc.)

Many dictionaries provide an explanation and/or a list of such usage labels in a preface or appendix. The type of the usage information may be indicated in the *type* attribute on the <usg> element. Some typical values are:

**geo** geographic area

**time** temporal, historical era ('archaic', 'old', etc.)

**dom** domain

**reg** register

**style** style (figurative, literal, etc.)

**plev** preference level ('chiefly', 'usually', etc.)

**acc** acceptability

**lang** language for foreign words, spellings pronunciations, etc.

**gram** grammatical usage

In addition to this kind of information, multilingual dictionaries often provide 'semantic cues' to help the user determine the right sense of a word in the source language (and hence the correct translation). These include synonyms, concept subdivisions, typical subjects and objects, typical verb complements, etc. These labels may also be marked with the <usg> element; sample values for the *type* attribute in these cases include:

**syn** synonym given to show use

**hyper** hypernym given to show usage

**colloc** collocation given to show usage

**comp** typical complement

**obj** typical object

**subj** typical subject

**verb** typical verb

**hint** unclassifiable piece of information to guide sense choice

In this entry, one spelling is marked as geographically restricted:

> **colour** *or US* **color** ... CED

```
<form>
  <orth>colour</orth>
  <form>
    <usg  type="geo">US</usg>
    <orth>color</orth>
  </form>
</form>
```

In the next example, usage labels are used to indicate domains, register, and synonyms associated with different senses:

> **palette** [palEt] nf (a) (Peinture: lit, fig) palette. (b) (Boucherie) shoulder. (c) (aube de roue) paddle; (battoir à linge) beetle; (Manutention, Constr) pallet. CR

```
<sense  n="a">
  <usg  type="dom">Peinture</usg>
  <usg  type="style">lit</usg>
  <usg  type="style">fig</usg>
  <cit  type="translation"  xml:lang="en">
    <quote>palette</quote>
  </cit>
</sense>
<sense  n="b">
  <usg  type="dom">Boucherie</usg>
  <cit  type="translation"  xml:lang="en">
    <quote>shoulder</quote>
  </cit>
</sense>
<sense  n="c">
  <sense>
    <usg  type="syn">aube de roue</usg>
    <cit  type="translation"  xml:lang="en">
      <quote>paddle</quote>
    </cit>
  </sense>
  <sense>
    <usg  type="syn">battoir à linge</usg>
    <cit  type="translation"  xml:lang="en">
      <quote>beetle</quote>
    </cit>
  </sense>
  <sense>
    <usg  type="dom">Manutention</usg>
    <usg  type="dom">Constr</usg>
    <cit  type="translation"  xml:lang="en">
```

```
      <quote>pallet</quote>
    </cit>
  </sense>
</sense>
```

When the usage label is hard to classify, it may be described as a 'hint':

> **rempaillage** [...] nm reseating, rebottoming (*with straw*). CR

```
<entry>
  <cit type="translation" xml:lang="en">
    <quote>reseating</quote>
    <quote>rebottoming</quote>
    <usg type="hint">with straw</usg>
  </cit>
</entry>
```

#### 10.3.5.3 Cross-References to Other Entries

Dictionary entries frequently refer to information in other entries, often using extremely dense notations to convey the headword of the entry to be sought, the particular part of the entry being referred to, and the nature of the information to be sought there (synonyms, antonyms, usage notes, etymology, an illustration, etc.)

Cross-references may be tagged in dictionaries using the <ref> and <ptr> elements defined in the core module (section *3.7. Simple Links and Cross-References*). In addition, the <xr> element may be used to group all the information relating to a cross-reference.

`<xr>` (cross-reference phrase) contains a phrase, sentence, or icon referring the reader to some other location in this or another text.

`<ref>` (reference) defines a reference to another location, possibly modified by additional text or comment.

`<ptr>` (pointer) defines a pointer to another location.

`<lbl>` (label) contains a label for a form, example, translation, or other piece of information, e.g. abbreviation for, contraction of, literally, approximately, synonyms:, etc.

As in other types of text, the actual pointing element (e.g. <ref> or <ptr>) is used to tag the cross-reference target proper (in dictionaries, usually the headword, possibly accompanied by a homograph number, a sense number, or other further restriction specifying what portion of the target entry is being referred to). The <xr> element is used to group the target with any accompanying phrases or symbols used to label the cross-reference; the cross-reference label itself may be encoded with <lbl> or may remain untagged. Both of the following are thus legitimate:

> **glee** ... Compare **madrigal** (1) CED

```
<entry>
  <form>
    <orth>glee</orth>
  </form>
  <xr>Compare <ptr target="#madrigal.1"/>
  </xr>
</entry>
<entry xml:id="madrigal.1">
  <form>
<!-- ... -->
  </form>
</entry>
```

**hostellerie** Syn. de hôtellerie (sens 1). DNT

```
<xr type="syn">
  <lbl>Syn. de</lbl>
  <ref>hôtellerie (sens 1)</ref>.
</xr>
```

In addition to using, or not using, <lbl> to mark the cross-reference label, the two examples differ in another way. The former assumes that the first sense of *madrigal* has the identifier madrigal.1, and that the specific form of the reference in the source volume can be reconstructed, if needed, from that information. The latter does not require the first sense of 'hôtellerie' to have an identifier, and retains the print form of the cross-reference; by omitting the *target* attribute of the <ref> element, however, the second example does assume implicitly either that some software could usefully parse the phrase tagged as a <ref> and find the location referred to, or else that such processing will not be necessary.

The *type* attribute on the pointing element or on the <xr> element may be used to indicate what kind of cross-reference is being made, using any convenient typology. Since different dictionaries may label the same kind of cross-reference in different ways, it may be useful to give normalized indications in the *type* attribute, enabling the encoder to distinguish irregular forms of cross-reference more reliably:

**rose**[2] ... *vb* the past tense of **rise** CED

```
<entry n="2">
  <form>
    <orth>rose</orth>
  </form>
  <xr type="inflectedForm">
    <lbl>the past tense of</lbl>
    <ref target="#rise">rise</ref>
  </xr>
</entry>
<!-- ... -->
<entry xml:id="rise">
  <form>
    <orth>rise</orth>
  </form>
<!-- main entry for "rise" as verb -->
</entry>
```

from cross-references for synonyms and the like:

**antagonist** ... syn see **adverse** W7

```
<xr type="synonym">
  <lbl>syn see</lbl>
  <ref target="#adverse">adverse</ref>
</xr>
<!-- ... -->
<entry xml:id="adverse">
  <form>
    <orth>adverse</orth>
  </form>
<!-- list of synonyms for "adverse" -->
</entry>
```

Strictly speaking, the reference above is not to the entry for *adverse*, but to the list of synonyms found within that entry. In some cases, the cross-reference is to a particular subset of the meanings of the entry in question:

> **globe** ...V. **armillaire** (sphère) PR

```
<xr>V. <ref  target="#armillaire">armillaire</ref>
  <lbl  type="sense-restriction">sphère</lbl>
</xr>
```

Cross-references occasionally occur in definition texts, example texts, etc., or may be free-standing within an entry. These may typically be encoded using <ref> or <ptr>, without an enclosing <xr>. For example:

> **entacher** ... *Acte entaché de nullité*, contenant un vice de forme ou passé par un incapable*. DNT

The asterisk signals a reference to the entry for *incapable*.

```
<def>contenant un vice de forme ou passé par un <ptr  target="#incapable"/>.</def>
```

In some cases, the form in the definition is inflected, and thus <ref> must be used to indicate more exactly the intended target, as here:

> **justifier** ...4. IMPRIM Donner a (une ligne) une longeur convenable au moyen de blancs (2, sens 1, 3). DNT

```
<sense  n="4">
  <usg  type="dom">imprim</usg>
  <def>Donner a (une ligne) une longeur convenable au moyen de
  <ref  target="#blanc-2.1.3">blancs (2, sens 1, 3)</ref>
  </def>
</sense>
<entry  xml:id="blanc"  n="2">
<!-- ... -->
  <sense  n="1">
<!-- ... -->
    <def  xml:id="blanc-2.1.3">...</def>
<!-- ... -->
  </sense>
<!-- ... -->
</entry>
```

#### 10.3.5.4 Notes within Entries

Dictionaries may include extensive explanatory notes about usage, grammar, context, etc. within entries. Very often, such notes appear as a separate section at the end of an entry. The standard <note> element should be used for such material.

`<note>` (note) contains a note or annotation.

For example:

> **neither** ('naɪðə, 'ni:ðə) *determiner* **1a** not one nor the other (of two); not either: *neither foot is swollen* ... **usage** A verb following a compound subject that uses *neither...* should be in the singular if both subjects are in the singular: *neither Jack nor John has done the work* CED

```
<entry>
  <form  type="contraction">
    <orth>neither</orth>
    <pron  notation="ipa">'naɪðə</pron>,
  <pron  notation="ipa">'ni:ðə</pron>
  </form>
<!-- ... -->
  <cit  type="example">
    <quote>neither foot is swollen</quote>
  </cit>
  <note  type="usage">A verb following a compound subject
     that uses <hi  rend="italic">neither</hi>… should be
     in the singular if both subjects are in the singular:
  <hi  rend="italic">neither Jack nor John has done the work</hi>
  </note>
</entry>
```

The formal declaration for <note> is given in section *3.9. Notes, Annotation, and Indexing*.

### 10.3.6 Related Entries

The <re> element encloses a degenerate entry which appears in the body of another entry for some purpose. Many dictionaries include related entries for direct derivatives or inflected forms of the entry word, or for compound words, phrases, collocations, and idioms containing the entry word.

Related entries can be complex, and may in fact include any of the information to be found in a regular entry. Therefore, the <re> element is defined to contain the same elements as an <entry> element.

Examples:

> **bevvy** (ˈbɛvɪ) *informal* n, pl **-vies** **1** a drink, esp an alcoholic one: *we had a few bevvies last night* **2** a session of drinking. ▷ *vb* **-vies**, **-vying**, **-vied** (*intr*) **3** to drink alcohol [probably from Old French *bevee*, *buvee*, drinking] > ’**bevvied** *adj* CED

```
<entry>
  <form>
    <orth>bevvy</orth>
    <pron  notation="ipa">'bɛvɪ</pron>
  </form>
  <usg  type="reg">informal</usg>
  <hom>
    <gramGrp>
      <pos>n</pos>
    </gramGrp>
    <sense  n="1">
      <def>a drink, esp. an alcoholic one: we had a few bevvies last night.</def>
    </sense>
  </hom>
<!-- ... sense 2 ... -->
  <hom>
    <gramGrp>
      <pos>vb</pos>
    </gramGrp>
    <sense  n="3">
      <def>to drink alcohol</def>
    </sense>
  </hom>
  <etym>probably from <lang>Old French</lang>
```

```
    <mentioned>bevee</mentioned>, <mentioned>buvee</mentioned>
    <gloss>drinking</gloss>
  </etym>
  <entry  type="relatedEntry"
    subtype="derived">
    <form>
      <orth>bevvied</orth>
    </form>
    <gramGrp>
      <pos>adj</pos>
    </gramGrp>
  </entry>
</entry>
```

## 10.4 Headword and Pronunciation References

Examples, definitions, etymologies, and occasionally other elements such as cross-references, orthographic forms, etc., often contain a shortened or iconic reference to the headword, rather than repeating the headword itself. The references may be to the orthographic form or to the pronunciation, to the form given or to a variant of that form. The following elements are used to encode such iconic references to a headword:

`<oRef>` (orthographic-form reference) in a dictionary example, indicates a reference to the orthographic form(s) of the headword.

`@type` indicates the kind of typographic modification made to the headword in the reference.

`<pRef>` (pronunciation reference) in a dictionary example, indicates a reference to the pronunciation(s) of the headword.

These elements all inherit the following attributes from the class `att.pointing` which may optionally be used to resolve any ambiguity about the headword form being referred to.

`att.pointing` provides a set of attributes used by all elements which point to other elements by means of one or more URI references.

`@target` specifies the destination of the reference by supplying one or more URI References.

Headword references come in a variety of formats:

**~** indicates a reference to the full form of the headword

**pref~** gives a prefix to be affixed to the headword

**~suf** gives a suffix to be affixed to the headword

**A~** gives the first letter in uppercase, indicating that the headword is capitalized

**pref~suf** gives a prefix and a suffix to be affixed to the headword

**a.** gives the initial of the word followed by a full stop, to indicate reference to the full form of the headword

**A.** refers to a capitalized form of the headword

The <oRef> element should be used for iconic or shortened references to the orthographic form(s) of the headword itself. It is an empty element and replaces, rather than enclosing, the reference. Note that the reference to a headword is not necessarily a simple string replacement. In the example '**colour**1, (US = color) ...~films; ~TV; Red, blue and yellow are ~s.' OALD, the tilde stands for either headword form (*colour*, *color*).

Examples:

**colonel** ... army officer above a lieutenant-~. OALD

```
<def>army officer above a lieutenant-<oRef/>
</def>
```

**academy** … The Royal A~of Arts OALD

```
<q>The Royal <oRef  type="cap"/> of Arts</q>
```

The following example demonstrates the use of the *target* attribute to refer to a specific form of the headword:

**vag-** or **vago-** comb form … : vagus nerve < **vag**al > < **vago**tomy > W7

```
<entry>
  <form>
    <orth  xml:id="di-o1">vag-</orth>
    <orth  xml:id="di-o2">vago-</orth>
  </form>
  <def>vagus nerve</def>
  <cit  type="example">
    <quote>
      <oRef  target="#di-o1"  type="noHyph"/>al</quote>
    <quote>
      <oRef  target="#di-o2"  type="noHyph"/>tomy</quote>
  </cit>
</entry>
```

In many cases the reference is not to the orthographic form of the headword, but rather to another form of the headword—usually to an inflected form. In these cases, the element <oRef> should be used; this element may take as its content the string as it appears in the text.

**take** … < Mr Burton **took** us for French > NPEG

```
<cit  type="example">
  <quote>Mr Burton <oRef  type="pt">took</oRef> us for French</quote>
</cit>
```

**take** … < was quite **~n** with him > NPEG

```
<cit  type="example">
  <quote>was quite <oRef  type="pp">
     <oRef/>n</oRef> with him</quote>
</cit>
```

The next example shows a discontinuous reference, using the attributes *next* and *prev*, which are defined in the additional module for linking, segmentation, and alignment (see chapter *17. Linking, Segmentation, and Alignment*) and therefore require that that module be selected in addition to that for dictionaries.

**mix up**… < it's easy to **mix** her **up** with her sister > NPEG

```
<cit  type="example">
  <quote>it's easy to <oRef  next="#ov2"  xml:id="ov1">mix</oRef>
    her <oRef  prev="#ov1"  xml:id="ov2">up</oRef> with her sister</quote>
</cit>
```

In addition, some dictionaries make reference to the pronunciation of the headword in the pronunciation of related entries, variants, or examples. The <pRef> element should be used for such references.

**hors d'oeuvre** /,aw'duhv (Fr O:r dœvr)/ n, pl hors d'oeuvres also hors d'oeuvre /'duhv(z) (Fr ~)/ NPEG

```
<form>
  <orth>hors d'oeuvre</orth>
  <pron notation="ipa">,aw'duhv</pron>
  <form>
    <usg type="lang">Fr</usg>
    <pron notation="ipa" xml:id="di-p2">O:r dœvr</pron>
  </form>
</form>
<form type="inflected">
  <number>pl</number>
  <orth>hors d'oeuvres</orth>
  <orth>hors d'oeuvre</orth>
  <pron notation="ipa" extent="part">'duhv(z)</pron>
  <form>
    <usg type="lang">Fr</usg>
    <pron>
      <pRef target="#di-p2"/>
    </pron>
  </form>
</form>
```

Because headword and pronunciation references can occur virtually anywhere in an entry, the <oRef> and <pRef> elements may appear within any other element defined for dictionary entries.

Since existing printed dictionaries use different conventions for headword references (swung dash, first letter abbreviated form, capitalization, or italicization of the word, etc.) the exact method used should be documented in the header.

## 10.5 Typographic and Lexical Information in Dictionary Data

Among the many possible views of dictionaries, it is useful to distinguish at least the following three, which help to clarify some issues raised with particular urgency by dictionaries, on account of the complexity of both their typography and their information structure.

- (a) the *typographic view*—the two-dimensional printed page, including information about lineation, pagination, and other features of layout
- (b) the *editorial view*—the one-dimensional sequence of tokens which can be seen as the input to the typesetting process; the wording and punctuation of the text and the sequencing of items are visible in this view, but specifics of the typographic realization are not
- (c) the *lexical view*—this view includes the underlying information represented in a dictionary, without concern for its exact textual form

For example, a domain indication in a dictionary entry might be broken over a line and therefore hyphenated ('naut-' 'ical'); the typographic view of the dictionary preserves this information. In a purely editorial view, the particular form in which the domain name is given in the particular dictionary (as 'nautical', rather than 'naut.', 'Naut.', etc.) would be preserved, but the fact that the word was split across two lines with a soft hyphen would not. Font shifts might plausibly be included in either a strictly typographic or an editorial view. In the lexical view, the only information preserved concerning domain would be some standard symbol or string representing the nautical domain (e.g. 'naut.') regardless of the form in which it appears in the printed dictionary.

In practice, publishers begin with the lexical view—i.e., lexical data as it might appear in a database—and generate first the editorial view, which reflects editorial choices for a particular dictionary (such as the use of the abbreviation 'Naut.' for 'nautical', the fonts in which different types of information are to be rendered, etc.), and then the typographic view, which is tied to a specific printed rendering. Computational linguists and philologists often begin

with the typographic view and analyse it to obtain the editorial and/or lexical views. Some users may ultimately be concerned with retaining only the lexical view, or they may wish to preserve the typographic or editorial views as a reference text, perhaps as a guard against the loss or misinterpretation of information in the translation process. Some researchers may wish to retain all three views, and study their interrelations, since research questions may well span all three views.

In general, an electronic encoding of a text will allow the recovery of at least one view of that text (the one which guided the encoding); if editorial and typographic practices are consistently applied in the production of a printed dictionary, or if exceptions to the rules are consistently recorded in the electronic encoding, then it is *in principle* possible to recover the editorial view from an encoding of the lexical view, and the typographic view from an encoding of the editorial view. In practice, of course, the severe compression of information in dictionaries, the variety of methods by which this compression is achieved, the complexity of formulating completely explicit rules for editorial and typographic practice, and the relative rarity of complete consistency in the application of such rules, all make the mechanical transformation of information from one view into another something of a vexed question.

This section describes some principles which may be useful in capturing one or the other of these views as consistently and completely as possible, and describes some methods of attempting to capture more than one view in a single encoding. Only the editorial and lexical views are explicitly treated here; for methods of recording the physical or typographic details of a text, see chapter *12. Representation of Primary Sources*. Other approaches to these problems, such as the use of repetitive encoding and links to show their correspondences, or the use of feature structures to capture the information structure, and of the *ana* and *inst* attributes to link feature structures to a transcription of the editorial view of a dictionary, are not discussed here (for feature structures, see chapter *19. Feature Structures*. For linkage of textual form and underlying information, see chapter *18. Simple Analytic Mechanisms*).

### 10.5.1 Editorial View

Common practice in encoding texts of all sorts relies on principles such as the following, which can be used successfully to capture the editorial view when encoding a dictionary:

1. All characters of the source text should be retained, with the possible exception of *rendition text* (for which see further below).
2. Characters appearing in the source text should typically be given as character data content in the document, rather than as the value of an attribute; again, rendition text may optionally be excepted from this rule.
3. Apart from the characters or graphics in the source text, nothing else should appear as content in the document, although it may be given in attribute values.
4. The material in the source text should appear in the encoding in the same order. Complications of the character sequence by footnotes, marginal notes, etc., text wrapping around illustrations, etc., may be dealt with by the usual means (for notes, see section *3.9. Notes, Annotation, and Indexing*).[72]

In a very conservative transcription of the editorial view of a text, *rendition characters* (e.g. the commas, parentheses, etc., used in dictionary entries to signal boundaries among parts of the entry) and *rendition text* (for example, conjunctions joining alternate headwords, etc.) are typically retained. Removing the tags from such a transcription will leave all and only the characters of the source text, in their original sequence.[73]

Consider, for example, the following entry:

> **pinna** (ˈpɪnə) *n*, *pl* **-nae** (-ni:) *or* **-nas 1** any leaflet of a pinnate compound leaf *2 zoology* a feather, wing, fin, or similarly shaped part **3** another name for **auricle** (2). [C18: via New Latin from Latin: wing, feather, fin] CED

A conservative encoding of the editorial view of this entry, which retains all rendition text, might resemble the following:

[72] Complications of sequence caused by marginal or interlinear insertions and deletions, which are frequent in manuscripts, or by unconventional page layouts, as in concrete poetry, magazines with imaginative graphic designers, and texts about the nature of typography as a medium, typically do not occur in dictionaries, and so are not discussed here.

[73] This is a slight oversimplification. Even in conservative transcriptions, it is common to omit page numbers, signatures of gatherings, running titles and the like. The simple description above also elides, for the sake of simplicity, the difficulties of assigning a meaning to the phrase 'original sequence' when it is applied to the printed characters of a source text; the 'original sequence' retained or recovered from a conservative transcription of the editorial view is, of course, the one established during the transcription by the encoder.

```
<entry>
  <form>
    <orth>pinna</orth>
    <pron  notation="ipa">'pɪnə</pron>
  </form>
  <gramGrp>
    <pos>n</pos>, </gramGrp>
  <form  type="inflected">
    <number>pl</number>
    <form>
      <orth  type="lat"  extent="part">-nae</orth>
      <pron  extent="part">(-ni:)</pron>
    </form> or <orth  type="std"  extent="part">-nas</orth>
  </form>
  <sense  n="1">1 <def>any leaflet of a pinnate compound leaf</def>
  </sense>
  <sense  n="2">2 <usg  type="dom">zoology</usg>
    <def>a feather, wing, fin, or similarly shaped part</def>
  </sense>
  <sense  n="3">3 <xr  type="syn">
     <lbl>another name for</lbl>
     <ref  target="#auricle.2">auricle (2)</ref>
    </xr>
  </sense>
  <etym>[<date>C18</date>: via <lang>New Latin</lang> from <lang>Latin</lang>:
  <gloss>wing</gloss>, <gloss>feather</gloss>,
  <gloss>fin</gloss>]</etym>
</entry>
<entry  xml:id="auricle.2">
  <form>
<!-- ... -->
  </form>
</entry>
```

A somewhat simplified encoding of the editorial view of this entry might exploit the fact that rendition text is often systematically recoverable. For example, parentheses consistently appear around pronunciation in this dictionary, and thus are effectively implied by the start- and end-tags for <pron>.[74] In such an encoding, removing the tags should exactly reproduce the sequence of characters in the source, minus rendition text. The original character sequence can be recovered fully by replacing tags with any rendition text they imply.

Encoding in this way, the example given above might resemble the following. The <tagUsage> element in the header would be used to record the following patterns of rendition text:

- parentheses appear around <pron> elements
- commas appear before inflected forms
- the word 'or' appears before alternate forms
- brackets appear around the etymology
- full stops appear after <pos>, inflection information, and sense numbers
- senses are numbered in sequence unless otherwise specified using the global *n* attribute

```
<entry>
  <form>
```

[74] The omission of rendition text is particularly common in systems for document production; it is considered good practice there, since automatic generation of rendition text is more reliable and more consistent than attempting to maintain it manually in the electronic text.

```
    <orth>pinna</orth>
    <pron>"pIn@</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
  <form  type="inflected">
    <number>pl</number>
    <form>
      <orth  type="lat"  extent="part">-nae</orth>
      <pron  extent="part">-ni:</pron>
    </form>
    <orth  type="std"  extent="part">-nas</orth>
  </form>
  <sense  n="1">
    <def>any leaflet of a pinnate compound leaf.</def>
  </sense>
  <sense  n="2">
    <usg  type="dom">Zoology</usg>
    <def>a feather, wing, fin, or similarly shaped part.</def>
  </sense>
  <sense  n="3">
    <xr  type="syn">
      <lbl>another name for</lbl>
      <ref>auricle (sense 2).</ref>
    </xr>
  </sense>
  <etym>
    <date>C18</date>: via <lang>New Latin</lang> from <lang>Latin</lang>:
  <gloss>wing</gloss>, <gloss>feather</gloss>, <gloss>fin</gloss>
  </etym>
</entry>
```

When rendition text is omitted, it is recommended that the means to regenerate it be fully documented, using the <tagUsage> element of the TEI header.

If rendition text is used systematically in a dictionary, with only a few mistakes or exceptions, the global attribute *rend* may be used on any tag to flag exceptions to the normal treatment. The values of the *rend* attribute are not prescribed, but it can be used with values such as no-comma, no-left-paren, etc. Specific values can be documented using the <rendition> element in the TEI header.

In the following (imaginary) example, no left parenthesis precedes the pronunciation:

> **biryani** or **biriani** %bIrI"A:nI) any of a variety of Indian dishes … [from Urdu]

This irregularity can be recorded thus:

```
<entry>
  <form>
    <orth>biryani</orth>
    <orth>biriani</orth>
    <pron  rend="noleftparen">%bIrI"A:nI</pron>
  </form>
  <def>any of a variety of Indian dishes … </def>
  <etym>from <lang>Urdu</lang>
  </etym>
</entry>
```

### 10.5.2 Lexical View

If the text to be interchanged retains only the lexical view of the text, there may be no concern for the recoverability of the editorial (not to speak of the typographic) view of the text. However, it is strongly recommended that the TEI header be used to document fully the nature of all alterations to the original data, such as normalization of domain names, expansion of inflected forms, etc.

In an encoding of the lexical view of a text, there are degrees of departure from the original data: normalizing inconsistent forms like 'nautical', 'naut'., 'Naut.', etc., to 'nautical' is a relatively slight alteration; expansion of 'delay -ed -ing' to 'delay, delayed, delaying' is a more substantial departure. Still more severe is the rearranging of the order of information in entries; for example:

- reorganizing the order of elements in an entry to show their relationship, as in

  > **clem** (klɛm) *or* **clam** *vb* **clems**, **clemming**, **clemmed** *or* **clams**, **clamming**, **clammed** CED

  where in a strictly lexical view one might wish to group 'clem' and 'clam' with their respective inflected forms.
- splitting an entry into two separate entries, as in

  > **celi.bacy** /"selIb@sI/ n [U] state of living unmarried, esp as a religious obligation. celi.bate /"selIb@t/ n [C] unmarried person (esp a priest who has taken a vow not to marry). OALD

  For some purposes, this entry might usefully be split into an entry for 'celibacy' and a separate entry for 'celibate'.

An encoding which captures the lexical view of the example given in the previous section might look something like the following. In this encoding:

- abbreviated forms have been silently expanded
- some forms have been moved to allow related forms to be grouped together
- the part of speech information has been moved to allow all forms to be given together
- the cross-reference to 'auricle' has been simplified

```
<entry>
  <form>
    <orth>pinna</orth>
    <pron>"pIn@</pron>
    <form type="inflected">
      <number>pl</number>
      <form>
        <orth type="lat">pinnae</orth>
        <pron>'pIni:</pron>
      </form>
      <orth type="std">pinnas</orth>
    </form>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
  <sense n="1">
    <def>any leaflet of a pinnate compound leaf.</def>
  </sense>
  <sense n="2">
    <usg type="dom">Zoology</usg>
    <def>a feather, wing, fin, or similarly shaped part.</def>
  </sense>
  <sense n="3">
    <xr type="syn">
      <ptr target="#auricle.2"/>
    </xr>
  </sense>
```

```
  <etym>
    <date>C18</date>: via <lang>New Latin</lang> from <lang>Latin</lang>:
  <gloss>wing</gloss>, <gloss>feather</gloss>, <gloss>fin</gloss>
  </etym>
</entry>
```

Whether the given dictionary encoding focusses on the lexical view and thus approaches the status of lexical databases, or uses the typographic/editorial view approach and needs to communicate the sometimes informally stated values for the particular descriptive features, the issue of *interoperability* of the content and of the container objects becomes relevant, in view of the growing tendency to interlink pieces of information across Internet resources. In such cases, it becomes crucial to be able to encode the fact that whether the information on, for instance, the value of the grammatical category of Number is provided as 'sg.', 'sing.', 'Singular', or equivalently 'poj.' in Polish, or 'Ez.' in German, etc., what is actually referred to is always the same grammatical value that can be rendered with a plethora of markers, depending on the publisher, language, or lexicographic tradition. In order to signal that this variety of surface markers in fact indicate the same underlying value, it is possible to align them with an external inventory of standardized values. The TEI provides the `att.datcat` attribute class for the purpose of aligning grammatical (or indeed any sort of) categories as well as their values with a reference taxonomy of shared data categories.

In the example below, a fragment of the entry for *isotope* cited in section *10.3.2. Grammatical Information* is adorned by references to standardized definitions for *part of speech* (*datcat*) and *adjective* (*valueDatcat*). Depending on the status and extent of the dictionary, various strategies may be used to reduce the redundancy of references.

```
<entry>
<!--...-->
  <form>
    <orth>isotope</orth>
  </form>
  <gramGrp>
    <pos  datcat="http://hdl.handle.net/11459/CCR_C-396_5a972b93-2294-ab5c-a541-7c344c5f26c3"
      valueDatcat="http://hdl.handle.net/11459/CCR_C-1230_23653c21-fca1-edf8-fd7c-3df2d6499157">ad
  </gramGrp>
<!--...-->
</entry>
```

In the above example, alignment is performed against the CLARIN Concept Registry.

## 10.5.3 Retaining Both Views

It is sometimes desirable to retain both the lexical and the editorial view, in which case a potential conflict exists between the two. When there is a conflict between the encodings for the lexical and editorial views, the principles described in the following sections may be applied.

### 10.5.3.1 Using Attribute Values to Capture Alternate Views

If the order of the data is the same in both views, then both views may be captured by encoding one 'dominant' view in the character data content of the document, and encoding the other using attribute values on the appropriate elements. If all tags were to be removed, the remaining characters would be those of the dominant view of the text.

The attribute class `att.lexicographic` (which includes the attributes *norm* and *org* from class `att.lexicographic.normalized`) is used to provide attributes for use in encoding multiple views of the same dictionary entry. These attributes are available for use on all elements defined in this chapter when the base module for dictionaries is selected.

When the editorial view is dominant, the following attributes may be used to capture the lexical view:

`att.lexicographic` provides a set of attributes for specifying standard and normalized values, grammatical functions, alternate or equivalent forms, and information about composite parts.

@*norm [att.lexicographic.normalized]* (normalized) provides the normalized/standardized form of information present in the source text in a non-normalized form.

@*split* (split) gives the list of split values for a merged form.

When the lexical view is dominant, the following attributes may be used to record the editorial view:

att.lexicographic provides a set of attributes for specifying standard and normalized values, grammatical functions, alternate or equivalent forms, and information about composite parts.

@*orig [att.lexicographic.normalized]* (original) gives the original string or is the empty string when the element does not appear in the source text.

@*mergedIn* (merged into) gives a reference to another element, where the original appears as a merged form.

One attribute is useful in either view:

att.lexicographic provides a set of attributes for specifying standard and normalized values, grammatical functions, alternate or equivalent forms, and information about composite parts.

@*opt* (optional) indicates whether the element is optional or not.

For example, if the source text had the domain label 'naut.', it might be encoded as follows. With the editorial view dominant:

```
<usg norm="nautical" type="dom">naut.</usg>
```

The lexical view of the same label would transcribe the normalized form as content of the <usg> element, the typographic form as an attribute value:

```
<usg orig="naut." type="dom">nautical</usg>
```

If the source text gives inflectional information for the verb *delay* as 'delay, -ed, -ing', it might usefully be expanded to 'delayed, delayed, delaying'. An encoding of the editorial view might take this form:

```
<form>
  <orth>delay</orth>
  <form type="inflected">
    <orth norm="delayed" extent="part">-ed</orth>
    <tns norm="pst,pstp"/>
  </form>
  <form type="inflected">
    <orth norm="delaying" extent="part">-ing</orth>
    <tns norm="prsp"/>
  </form>
</form>
```

Note the use of the <tns> tag with null content, to enable the representation of implicit information even though it has no print realization.

The lexical view might be encoded thus:

```
<form>
  <orth>delay</orth>
  <form type="inflected">
    <orth orig="-ed">delayed</orth>
    <tns orig="">pst</tns>
```

```
    <tns  orig="">pstp</tns>
  </form>
  <form  type="inflected">
    <orth  orig="-ing">delaying</orth>
    <tns  orig="">prsp</tns>
  </form>
</form>
```

A particular problem may be posed by the common practice of presenting two alternate forms of a word in a single string, by marking some parts of the word as optional in some forms. The following entry is for a word which can be spelled either 'thyrostimuline' or 'thyréostimuline':

**thyr(é)ostimuline** [tiR(e)ostimylin] ...

With the editorial view dominant, this entry might begin thus:

```
<form>
  <orth  split="thyrostimuline, thyréostimuline">thyr(é)ostimuline</orth>
  <pron  split="tiRostimylin, tiReostimylin">tiR(e)ostimylin</pron>
</form>
```

With the lexical view dominant, however, two <orth> and two <pron> elements would be encoded, in order to disentangle the two forms; the *orig* attribute would be used to record the typographic presentation of the information in the source.

```
<form>
  <orth  xml:id="dic-o1"
    orig="thyr(é)ostimuline">thyrostimuline</orth>
  <pron  xml:id="dic-p1"
    orig="tiR(e)ostimylin">tiRostimylin</pron>
</form>
<form>
  <orth  mergedIn="#dic-o1">thyréostimuline</orth>
  <pron  mergedIn="#dic-p1">tiReostimylin</pron>
</form>
```

This example might also be encoded using the *opt* attribute combined with the attributes *next* and *prev* defined in chapter *17. Linking, Segmentation, and Alignment.*

```
<form>
  <orth  next="#dict-o2"  xml:id="dict-o1">thyr</orth>
  <orth  next="#dict-o3"  prev="#dict-o1"
    xml:id="dict-o2"  opt="true">é</orth>
  <orth  prev="#dict-o2"  xml:id="dict-o3">ostimuline</orth>
  <pron  next="#dict-p2"  xml:id="dict-p1">tiR</pron>
  <pron  next="#dict-p3"  prev="#dict-p1"
    xml:id="dict-p2"  opt="true">e</pron>
  <pron  prev="#dict-p2"  xml:id="dict-p3">ostimylin</pron>
</form>
```

Note that this transcription preserves both the lexical and editorial views in a single encoding. However, it has the disadvantage that the strings corresponding to entire words do not appear in the encoding uninterrupted, and therefore complex processing is required to retrieve them from the encoded text. The use of the *opt* attribute is recommended, however, when long spans of text are involved, or when the optional part contains embedded tags.

For example, the following gives two definitions in one text: ‘picture drawn with coloured chalk made into crayons’, and ‘coloured chalk made into crayons’:

> **pas.tel** /”pastl US: pa”stel/ n 1 (picture drawn with) coloured chalk made into crayons. 2… OALD

A simple encoding solution would be to leave the definition text unanalysed, but this might be felt inadequate since it does not show that there are two definitions. A possible alternative encoding would be:

```
<sense n="1">
  <def>coloured
     chalk made into crayons</def>
  <def>picture drawn with coloured chalk
     made into crayons</def>
</sense>
```

This transcribes some characters of the source text twice, however, which deviates from the usual practice. The following encoding records both the editorial and lexical views:

```
<sense n="1">
  <def next="#d2" xml:id="d1" opt="true">picture drawn
     with</def>
  <def prev="#d1" xml:id="d2">coloured chalk made into
     crayons</def>
</sense>
```

#### 10.5.3.2 Recording Original Locations of Transposed Elements

The attributes described in the previous section are useful only when the order of material is the same in both the editorial and the lexical view. When the two views impose different orders on the data, the standard linking mechanisms may be used to show the original location of material transposed in an encoding of the lexical view.

If the original is only slightly modified, the <anchor> element may be used to mark the original location of the material, and the *location* attribute may be used on the lexical encoding of that material to indicate its original location(s). Like those in the preceding section, this attribute is defined for the attribute class `att.lexicographic`:

`att.lexicographic` provides a set of attributes for specifying standard and normalized values, grammatical functions, alternate or equivalent forms, and information about composite parts.
`@opt` (optional) indicates whether the element is optional or not.

For example:

> **pinna** ('pɪnə) *n*, *pl* **-nae** (-ni:) *or* **-nas** CED

```
<form>
  <orth>pinna</orth>
  <pron notation="ipa">'pɪnə</pron>
  <anchor xml:id="p01"/>
  <form type="inflected">
    <number>pl</number>
    <form>
      <orth extent="part">-nae</orth>
      <pron extent="part">-ni:</pron>
    </form>
    <orth extent="part">-nas</orth>
  </form>
</form>
```

```
<gramGrp>
  <pos  location="#p01">n</pos>
</gramGrp>
```

## 10.6 Unstructured Entries

The content model for the <entry> element provides an entry structure suitable for many average dictionaries, as well as many regular entries in more exotic dictionaries. However, the structure of some dictionaries does not allow the restrictions imposed by the content model for <entry>. To handle these cases, the <entryFree> and <dictScrap> elements are provided to support much wider variation in entry structure. The <dictScrap> element offers less freedom, in that it can only contain phrase level elements, but it can itself appear at any point within a dictionary entry where any of the structural components of a dictionary entry are permitted. As such, it acts as a container for otherwise anomalous parts of an entry.

The <entryFree> element places no constraints at all upon the entry: any element defined in this chapter, as well as all the normal phrase-level and inter-level elements, can appear anywhere within it. With the <entryFree> element, the encoder is free to use any element anywhere, as well as to use or omit grouping elements such as <form>, <gramGrp>, etc.

The <entryFree> element allows the encoding of entries which violate the structure specified for the <entry> element. For example, in the following entry from a dictionary already in electronic form, it is necessary to include a <pron> element within a <def>. This is not permitted in the content model for <entry>, but it poses no problem in the <entryFree> element.

```
  <entb h="demigod"> <hwd>demi|god</hwd> <pr> <ph>"demIgQd</ph> </pr> <hps ps="n">
<hsn> <def>one who is partly divine and partly human</def> <def>(in Gk myth, etc) the
son of a god and a mortal woman, eg<cf>Hercules</cf> <pr> <ph>"h3:kjUli:z</ph> </pr>
</def> </hsn> </hps> </ent>
```

```
<entryFree>
  <form>
    <orth>demigod</orth>
    <hyph>demi|god</hyph>
    <pron>"demIgQd</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
  <def>one who is partly divine and partly human</def>
  <def>(in Gk myth, etc) the son of a god and a mortal woman, eg
  <mentioned>Hercules</mentioned>
  </def>
  <pron>"h3:kjUli:z</pron>
</entryFree>
```

The <entryFree> element also makes it possible to transcribe a dictionary using only phrase-level ('atomic') elements—that is, using no grouping elements at all. This can be desirable if the encoder wants a completely 'flat' view, with no indication of or commitment to the association of one element with another. The following encoding uses no grouping elements, and keeps all rendition text:

**biryani** *or* **biriani** (ˌbɪrɪˈɑːnɪ) *n* any of a variety of Indian dishes ... [from Urdu] CED

```
<entryFree>
  <orth>biryani</orth> or <orth>biriani</orth>
  <pron  notation="ipa">(ˌbɪrɪ'a:nɪ)</pron>
  <def>any of a variety of Indian dishes …</def>
  <etym>[from <lang>Urdu</lang>]</etym>
</entryFree>
```

Here is an alternative way of representing the same structure, this time using <dictScrap>:

```
<entry>
  <dictScrap>
    <orth>biryani</orth> or <orth>biriani</orth>
    <pron  notation="ipa">(ˌbɪrɪ'a:nɪ)</pron>
    <def>any of a variety of Indian dishes …</def>
    <etym>[from <lang>Urdu</lang>]</etym>
  </dictScrap>
</entry>
```

## 10.7 The Dictionary Module

The module defined in this chapter makes available the following components:

**Module dictionaries: Dictionaries**

- *Elements defined*: case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number oRef orth pRef per pos pron re sense stress subc superEntry syll tns usg xr
- *Classes defined*: att.entryLike att.lexicographic model.entryLike model.formPart model.gramPart model.lexicalRefinement model.morphLike model.ptrLike.form

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 11 Manuscript Description

## 11.1 Overview

The `msdescription` module[75] defines a special purpose element which can be used to provide detailed descriptive information about handwritten primary sources and other text-bearing objects. Although originally developed to meet the needs of cataloguers and scholars working with medieval manuscripts in the European tradition, the scheme presented here is general enough that it can also be extended to other traditions and materials, and is potentially useful for any kind of text-bearing artefact. Where the textuality of an object is not the primary concern, encoders may wish to use the <object> element which provides a very similar system of description (see *14.3.6. Objects*.

The scheme described here is also intended to accommodate the needs of many different classes of encoders. On the one hand, encoders may be engaged in *retrospective conversion* of existing detailed descriptions and catalogues into machine tractable form; on the other, they may be engaged in cataloguing *ex nihilo*, that is, creating new detailed descriptions for materials never before catalogued. Some may be primarily concerned to represent accurately the description itself, as opposed to the ideas and interpretations the description represents; others may have entirely opposite priorities. At one extreme, a project may simply wish to capture an existing catalogue in a form that can be displayed on the Web, and which can be searched for literal strings, or for such features such as titles, authors and dates; at the other, a project may wish to create, in highly structured and encoded form, a detailed database of information about the physical characteristics, history, interpretation, etc. of the material, able to support practitioners of *quantitative codicology* as well as librarians.

To cater for this diversity, here as elsewhere, these Guidelines propose a flexible strategy, in which encoders must choose for themselves the approach appropriate to their needs, and are provided with a choice of encoding mechanisms to support those differing degrees.

## 11.2 The Manuscript Description Element

The <msDesc> element will normally appear within the <sourceDesc> element of the header of a TEI-conformant document, where the document being encoded is a digital representation of some manuscript original, whether as an encoded transcription, as a collection of digital images (as described in *12.1. Digital Facsimiles*), or as some combination of the two. However, in cases where the document being encoded is essentially a collection of manuscript descriptions, the <msDesc> element may be used in the same way as the bibliographic elements (<bibl>, <biblFull>, and <biblStruct>) making up the TEI element class `model.biblLike`. These typically appear within the <listBibl> element.

`<msDesc>` (manuscript description) contains a description of a single identifiable manuscript or other text-bearing object such as an early printed book.

The <msDesc> element has the following components, which provide more detailed information under a number of headings. Each of these component elements is further described in the remainder of this chapter.

`<msIdentifier>` (manuscript identifier) contains the information required to identify the manuscript or similar object being described.

`<head>` (heading) contains any type of heading, for example the title of a section, or the heading of a list, glossary, manuscript description, etc.

`<msContents>` (manuscript contents) describes the intellectual content of a manuscript, manuscript part, or other object either as a series of paragraphs or as a series of structured manuscript items.

`<physDesc>` (physical description) contains a full physical description of a manuscript, manuscript part, or other object optionally subdivided using more specialized elements from the `model.physDescPart` class.

`<history>` (history) groups elements describing the full history of a manuscript, manuscript part, or other object.

`<additional>` (additional) groups additional information, combining bibliographic information about a manuscript or other object, or surrogate copies of it, with curatorial or administrative information.

[75] This chapter is based on the work of the European MASTER (Manuscript Access through Standards for Electronic Records) project, funded by the European Union from January 1999 to June 2001, and led by Peter Robinson, then at the Centre for Technology and the Arts at De Montfort University, Leicester (UK). Significant input also came from a TEI Workgroup headed by Consuelo W. Dutschke of the Rare Book and Manuscript Library, Columbia University (USA) and Ambrogio Piazzoni of the Biblioteca Apostolica Vaticana (IT) during 1998-2000.

`<msPart>` (manuscript part) contains information about an originally distinct manuscript or part of a manuscript, which is now part of a composite manuscript.

`<msFrag>` (manuscript fragment) contains information about a fragment described in relation to a prior context, typically as a description of a virtual reconstruction of a manuscript or other object whose fragments were catalogued separately.

The first of these components, <msIdentifier>, is the only one which is mandatory; it is described in more detail in *11.4. The Manuscript Identifier* below. It is followed optionally by one or more <head> elements, each holding a brief heading (see *11.5. The Manuscript Heading*), and then either one or more paragraphs, marked up as a series of <p> elements, or at most one of each of the specialized elements <msContents> (*11.6. Intellectual Content*), <physDesc> (*11.7. Physical Description*), <history> (*11.8. History*), and <additional> (*11.9. Additional Information*). These elements are all optional, and if used they may appear at most once in a given <msDesc>, in any order. Finally, in the case of a composite manuscript (a manuscript composed of several codicological units) or a fragmented manuscript (a manuscript whose parts are now dispersed and kept at different places), a full description may also contain one or more <msPart> (*11.10. Manuscript Parts*) elements and <msFrag> (*11.11. Manuscript Fragments*) elements, respectively.

To demonstrate the use of this module, consider the following sample manuscript description, chosen more or less at random from the Bodleian Library's *Summary catalogue* ([291])

**28843.** In Latin, on parchment: written in more than one hand of the 13th cent. in England: $7\frac{1}{4} \times 5\frac{3}{8}$ in., i + 55 leaves, in double columns: with a few coloured capitals.

'Hic incipit Bruitus Anglie,' the De origine et gestis Regum Angliae of Geoffrey of Monmouth (Galfridus Monumetensis): *beg.* 'Cum mecum multa & de multis.'

On fol. 54ᵛ very faint is 'Iste liber est fratris guillelmi de buria de . . . Roberti ordinis fratrum Pred[icatorum],' 14th cent. (?): 'hanauilla' is written at the foot of the page (15th cent.). Bought from the rev. W. D. Macray on March 17, 1863, for £1 10*s*.

Now MS. **Add. A. 61.**

Figure 13: Entry for Bodleian MS. Add. A. 61 in Madan et al. 1895-1953

The simplest way of digitizing this catalogue entry would simply be to key in the text, tagging the relevant parts of it which make up the mandatory <msIdentifier> element, as follows:

```
<msDesc>
  <msIdentifier>
    <settlement>Oxford</settlement>
    <repository>Bodleian Library</repository>
    <idno>MS. Add. A. 61</idno>
    <altIdentifier  type="SC">
      <idno>28843</idno>
    </altIdentifier>
  </msIdentifier>
  <p>In Latin, on parchment: written in more than one hand of the 13th cent. in
     England: 7¼ x 5⅜ in., i + 55 leaves, in double columns: with a few coloured
     capitals.</p>
  <p>'Hic incipit Bruitus Anglie,' the De origine et gestis Regum Angliae of
     Geoffrey of Monmouth (Galfridus Monumetensis: beg. 'Cum mecum multa & de
     multis.'</p>
```

```
  <p>On fol. 54v very faint is 'Iste liber est fratris guillelmi de buria de ...
     Roberti ordinis fratrum Pred[icatorum],' 14th cent. (?): 'hanauilla' is written
     at the foot of the page (15th cent.). Bought from the rev. W. D. Macray on March
     17, 1863, for £1 10s.</p>
</msDesc>
```

With a suitable stylesheet, this encoding would be as readable as the original; it would not, however, be very useful for search purposes since only shelfmarks and other identifiers are distinguished. To improve on this, one might wrap the paragraphs in the appropriate special-purpose first-child-level elements of <msDesc> and add some of the additional phrase-level elements available when this module is in use:

```
<msDesc>
  <msIdentifier>
    <settlement>Oxford</settlement>
    <repository>Bodleian Library</repository>
    <idno>MS. Add. A. 61</idno>
    <altIdentifier  type="SC">
      <idno>28843</idno>
    </altIdentifier>
  </msIdentifier>
  <msContents>
    <p>
      <quote>Hic incipit Bruitus Anglie,</quote> the <title>De origine et gestis
           Regum Angliae</title> of Geoffrey of Monmouth (Galfridus Monumetensis):
beg.
    <quote>Cum mecum multa & de multis.</quote> In Latin.</p>
  </msContents>
  <physDesc>
    <p>
      <material  function="support">Parchment</material>: written in
        more than one hand: 7¼ x 5⅜ in., i + 55 leaves, in double columns:
        with a few coloured capitals.</p>
  </physDesc>
  <history>
    <p>Written in <origPlace>England</origPlace> in the <origDate>13th
           cent.</origDate> On fol. 54v very faint is <quote>Iste liber est fratris
           guillelmi de buria de ... Roberti ordinis fratrum Pred[icatorum],</quote>
14th
        cent. (?): <quote>hanauilla</quote> is written at the foot of the page (15th
        cent.). Bought from the rev. W. D. Macray on March 17, 1863, for £1 10s.</p>
  </history>
</msDesc>
```

Note that in this version the text has been slightly reorganized, but no actual rewriting has been necessary. The encoding now allows the user to search for such features as title, material, and date and place of origin; it is also possible to distinguish quoted material from descriptive passages and to search within descriptions relating to a particular topic (for example, history as distinct from material).

This process could be continued further, restructuring the whole entry so as to take full advantage of many more of the encoding possibilities provided by the module described in this chapter:

```
<msDesc>
  <msIdentifier>
    <settlement>Oxford</settlement>
    <repository>Bodleian Library</repository>
```

```
    <idno>MS. Add. A. 61</idno>
    <altIdentifier type="SC">
      <idno>28843</idno>
    </altIdentifier>
  </msIdentifier>
  <msContents>
    <msItem>
      <author xml:lang="en">Geoffrey of Monmouth</author>
      <author xml:lang="la">Galfridus Monumetensis</author>
      <title type="uniform" xml:lang="la">De origine et gestis Regum Angliae</title>
      <rubric xml:lang="la">Hic incipit Bruitus Anglie</rubric>
      <incipit xml:lang="la">Cum mecum multa & de multis</incipit>
      <textLang mainLang="la">Latin</textLang>
    </msItem>
  </msContents>
  <physDesc>
    <objectDesc form="codex">
      <supportDesc material="perg">
        <support>
          <p>Parchment.</p>
        </support>
        <extent>i + 55 leaves <dimensions scope="all" type="leaf"
            unit="inch">
            <height>7¼</height>
            <width>5⅜</width>
          </dimensions>
        </extent>
      </supportDesc>
      <layoutDesc>
        <layout columns="2">
          <p>In double columns.</p>
        </layout>
      </layoutDesc>
    </objectDesc>
    <handDesc>
      <p>Written in more than one hand.</p>
    </handDesc>
    <decoDesc>
      <p>With a few coloured capitals.</p>
    </decoDesc>
  </physDesc>
  <history>
    <origin>
      <p>Written in <origPlace>England</origPlace> in the <origDate notAfter="1300"
          notBefore="1200">13th cent.</origDate>
      </p>
    </origin>
    <provenance>
      <p>On fol. 54v very faint is <quote xml:lang="la">Iste liber est fratris
              guillelmi de buria de <gap/> Roberti ordinis fratrum
              Pred<ex>icatorum</ex>
        </quote>, 14th cent. (?): <quote>hanauilla</quote> is
            written at the foot of the page (15th cent.).</p>
    </provenance>
    <acquisition>
      <p>Bought from the rev. <name key="MCRAYWD">W. D. Macray</name> on
<date when="1863-03-17">March 17, 1863</date>, for £1 10s.</p>
```

```
    </acquisition>
  </history>
</msDesc>
```

In the remainder of this chapter we discuss all of the encoding features demonstrated above, together with many other related matters.

## 11.3 Phrase-level Elements

When the `msdescription` module is in use, several extra elements are added to the phrase level class, and thus become available within paragraphs and elsewhere in the document. These elements are listed below in alphabetical order:

`<catchwords>` (catchwords) describes the system used to ensure correct ordering of the quires or similar making up a codex, incunable, or other object typically by means of annotations at the foot of the page.

`<dimensions>` (dimensions) contains a dimensional specification.

`<heraldry>` (heraldry) contains a heraldic formula or phrase, typically found as part of a blazon, coat of arms, etc.

`<locus>` (locus) defines a location within a manuscript, manuscript part, or other object typically as a (possibly discontinuous) sequence of folio references.

`<locusGrp>` (locus group) groups a number of locations which together form a distinct but discontinuous item within a manuscript, manuscript part, or other object.

`<material>` (material) contains a word or phrase describing the material of which the object being described is composed.

`<watermark>` (watermark) contains a word or phrase describing a watermark or similar device.

`<objectType>` (object type) contains a word or phrase describing the type of object being referred to.

`<origDate>` (origin date) contains any form of date, used to identify the date of origin for a manuscript, manuscript part, or other object.

`<origPlace>` (origin place) contains any form of place name, used to identify the place of origin for a manuscript, manuscript part, or other object.

`<secFol>` (second folio) marks the word or words taken from a fixed point in a codex (typically the beginning of the second leaf) in order to provide a unique identifier for it.

`<signatures>` (signatures) contains discussion of the leaf or quire signatures found within a codex or similar object.

Within a manuscript description, many other standard TEI phrase level elements are available, notably those described in the Core module (*3. Elements Available in All TEI Documents*). Additional elements of particular relevance to manuscript description, such as those for names and dates, may also be made available by including the relevant module in one's schema.

### 11.3.1 Origination

The following elements may be used to provide information about the origins of any aspect of a manuscript:

`<origDate>` (origin date) contains any form of date, used to identify the date of origin for a manuscript, manuscript part, or other object.

`<origPlace>` (origin place) contains any form of place name, used to identify the place of origin for a manuscript, manuscript part, or other object.

The <origDate> and <origPlace> elements are specialized forms of the existing <date> and <name> elements respectively, used to indicate specifically the date and place of origin of a manuscript or manuscript part. Such information would normally be encoded within the <history> element, discussed in section *11.8. History*. <origDate> and <origPlace> can also be used to identify the place or date of origin of any aspect of the manuscript, such as its decoration or binding, when these are not of the same date or from the same location as rest of the manuscript. Both these elements are members of the `att.editLike` class, from which they inherit many attributes.

The <origDate> element is a member of the `att.datable` class, and may thus also carry additional attributes giving normalized values for the associated dating.

### 11.3.2 Material and Object Type

The <material> element can be used to tag any specific term used for the physical material of which a manuscript (or binding, seal, etc.) is composed. The <objectType> element may be used to tag any term specifying the type of object or manuscript upon with the text is written.

`<material>` (material) contains a word or phrase describing the material of which the object being described is composed.

`@function` describes the function or use of the material in relation to the object as a whole.

`<objectType>` (object type) contains a word or phrase describing the type of object being referred to.

These elements may appear wherever a term regarded as significant by the encoder occurs, as in the following examples:

```
<binding>
  <p>Brown <material>calfskin</material>, previously with two clasps.</p>
</binding>
```

```
<support>
  <p>
    <material  function="support">Parchment</material>
    <objectType>codex</objectType> with half <material  function="binding">goat-leather</material>
     binding.</p>
</support>
```

### 11.3.3 Watermarks and Stamps

Two further elements are provided to mark up other decorative features characteristic of manuscript leaves and bindings:

`<watermark>` (watermark) contains a word or phrase describing a watermark or similar device.

`<stamp>` (stamp) contains a word or phrase describing a stamp or similar device.

These elements may appear wherever a term regarded as significant by the encoder occurs. The <watermark> element is most likely to be of use within the <support> element discussed in *11.7.1.1. Support* below. We give a simple example here:

```
<support>
  <material>Rag
     paper</material> with <watermark>anchor</watermark> watermark
</support>
```

The <stamp> element will typically appear when text from the source is being transcribed, for example within a rubric in the following case:

```
<rubric>
  <lb/>Apologyticu TTVLLIANI AC IGNORATIA IN XPO IHV
<lb/>SI NON LICET
<lb/>NOBIS RO
<lb/>manii imperii <stamp>Bodleian stamp</stamp>
  <lb/>
</rubric>
```

It may also appear as part of the detailed description of a binding:

```
<binding>
  <p>Modern calf recasing with original armorial stamp <stamp>with legend
    <mentioned  xml:lang="la">Ex Bibliotheca J. Richard
           D.M.</mentioned>
    </stamp>
  </p>
</binding>
```

If, as here, any text contained by a stamp is included in its description it should be clearly distinguished from that description. The element <mentioned> may be used for this purpose, as shown above.

### 11.3.4 Dimensions

The <dimensions> element can be used to specify the size of some aspect of the manuscript, and thus may be thought of as a specialized form of the existing TEI <measure> element.

`<dimensions>` (dimensions) contains a dimensional specification.

- `@type` indicates which aspect of the object is being measured.

The <dimensions> element will normally occur within the element describing the particular feature or aspect of a manuscript whose dimensions are being given; thus the size of the leaves would be specified within the <support> or <extent> element (part of the <physDesc> element discussed in *11.7.1. Object Description*), while the dimensions of other specific parts of a manuscript, such as accompanying materials, binding, etc., would be given in other parts of the description, as appropriate.

The following elements are available within the <dimensions> element:

`<height>` (height) contains a measurement measured along the axis at a right angle to the bottom of the object.

`<width>` (width) contains a measurement of an object along the axis parallel to its bottom, e.g. perpendicular to the spine of a book or codex.

`<depth>` (depth) contains a measurement from the front to the back of an object, perpendicular to the measurement given by the <width> element.

`<dim>` contains any single measurement forming part of a dimensional specification of some sort.

These elements, as well as <dimensions> itself, are all members of the `att.dimensions` class, which also inherits attributes from the `att.ranging` class. They all thus carry the following attributes:

`att.dimensions` provides attributes for describing the size of physical objects.

- `@scope` where the measurement summarizes more than one observation, specifies the applicability of this measurement.
- `@extent` indicates the size of the object concerned using a project-specific vocabulary combining quantity and units in a single string of words.
- `@unit` names the unit used for the measurement
- `@quantity` specifies the length in the units specified

`att.ranging` provides attributes for describing numerical ranges.

- `@atLeast` gives a minimum estimated value for the approximate measurement.
- `@atMost` gives a maximum estimated value for the approximate measurement.
- `@min` where the measurement summarizes more than one observation or a range, supplies the minimum value observed.
- `@max` where the measurement summarizes more than one observation or a range, supplies the maximum value observed.

Attributes *scope*, *min*, and *max* are used only when the measurement applies to several items, for example the size of all leaves in a manuscript; attributes *atLeast* and *atMost* are used when the measurement applies to a single item, for example the size of a specific codex, but has had to be estimated. Attribute *quantity* is used when the measurement can be given exactly, and applies to a single item; this is the usual situation. In this case, the units in which dimensions are measured may be specified using the *unit* attribute, the value of which will normally be

taken from a closed set of values appropriate to the project, using standard units of measurement wherever possible, such as cm, mm, in, line, char. If however the only data available for the measurement uses some other unit, or it is preferred to normalize it in some other way, then it may be supplied as a string value by means of the *extent* attribute.

In the simplest case, only the *extent* attribute may be supplied:

```
<width extent="6 cubit">six
cubits</width>
```

More usually, the measurement will be normalized into a value and an appropriate SI unit:

```
<width quantity="270" unit="cm">six cubits</width>
```

Where the exact value is uncertain, the attributes *atLeast* and *atMost* may be used to indicate the upper and lower bounds of an estimated value:

```
<width atLeast="250" atMost="300" unit="cm">six cubits</width>
```

It is often convenient to supply a measurement which applies to a number of discrete observations: for example, the number of ruled lines on the pages of a manuscript (which may not all be the same), or the diameter of an object like a bell, which will differ depending where it is measured. In such cases, the *scope* attribute may be used to specify the observations for which this measurement is applicable:

```
<height unit="line" scope="most"
 atLeast="20"/>
```

This indicates that most pages have at least 20 lines. The attributes *min* and *max* can also be used to specify the possible range of values: for example, to show that all pages have between 12 and 30 lines:

```
<height unit="line" scope="all" min="12"
 max="30"/>
```

The <dimensions> element may be repeated as often as necessary, with appropriate attribute values to indicate the nature and scope of the measurement concerned. For example, in the following case the leaf size and ruled space of the leaves of the manuscript are specified:

```
<dimensions type="ruled" unit="mm">
 <height scope="most" quantity="90"
  unit="mm"/>
 <width scope="most" quantity="48" unit="mm"/>
</dimensions>
<dimensions type="leaves">
 <height min="157" max="160" unit="mm"/>
 <width quantity="105"/>
</dimensions>
```

This indicates that for most leaves of the manuscript being described the ruled space is 90 mm high and 48 mm wide, while the leaves throughout are between 157 and 160 mm in height and 105 mm in width.

The <dim> element is provided for cases where some measurement other than height, width, or depth is required. Its *type* attribute is used to indicate the type of measurement involved:

```
<dimensions unit="cm">
  <dim type="circumference" quantity="48"
    unit="mm"/>
  <height quantity="90" unit="mm"/>
</dimensions>
```

The order in which components of the <dimensions> element may be supplied is not constrained.

### 11.3.5 References to Locations within a Manuscript

The <locus> and its grouping element <locusGrp> element are specialized forms of the <ref> element, used to indicate a location, or sequence of locations, within a manuscript.

`<locus>` (locus) defines a location within a manuscript, manuscript part, or other object typically as a (possibly discontinuous) sequence of folio references.

- `@from` (from) specifies the starting point of the location in a normalized form, typically a page number.
- `@to` (to) specifies the end-point of the location in a normalized form, typically as a page number.
- `@scheme` (scheme) identifies the foliation scheme in terms of which the location is being specified by pointing to some <foliation> element defining it, or to some other equivalent resource.

`<locusGrp>` (locus group) groups a number of locations which together form a distinct but discontinuous item within a manuscript, manuscript part, or other object.

- `@scheme` (scheme) identifies the foliation scheme in terms of which all the locations contained by the group are specified by pointing to some <foliation> element defining it, or to some other equivalent resource.

The <locus> element is used to reference a single location within a manuscript, typically to specify the location occupied by the element within which it appears. If, for example, it is used as the first component of an <msItem> or <msItemStruct> element, or of any of the more specific elements appearing within one (see further section *11.6. Intellectual Content* below) then it is understood to specify the location (or locations) of that item within the manuscript being described.

#### 11.3.5.1 Identifying a Location

A <locus> element can be used to identify any reference to one or more folios within a manuscript, wherever such a reference is appropriate. Locations are conventionally specified as a sequence of folio or page numbers, but may also be a discontinuous list, or a combination of the two. This specification should be given as the content of the <locus> element, using the conventions appropriate to the individual scholar or holding institution, as in the following example:

```
<msItem n="1">
  <locus>ff. 1-24r</locus>
  <title>Apocalypsis beati Ioannis Apostoli</title>
</msItem>
```

A normalized form of the location can also be supplied, using special purpose attributes on the <locus> element, as in the following revision of the above example:

```
<msItem n="1">
  <locus from="1r" to="24r">ff. 1-24r</locus>
  <title>Apocalypsis beati Ioannis Apostoli</title>
</msItem>
```

When the item concerned occupies a discontinuous sequence of pages, this may simply be indicated in the body of the <locus> element:

```
<msItem  n="1">
  <locus>ff. 1-12v, 18-24r</locus>
  <title>Apocalypsis beati Ioannis Apostoli</title>
</msItem>
```

Alternatively, if it is desired to indicate normalized values for each part of the sequence, a sequence of <locus> elements can be supplied, grouped within the <locusGrp> element:

```
<msItem  n="1">
  <locusGrp>
    <locus  from="1r"  to="12v">ff. 1-12v</locus>
    <locus  from="18"  to="24r">ff. 18-24r</locus>
  </locusGrp>
  <title>Apocalypsis beati Ioannis Apostoli</title>
</msItem>
```

If an existing catalogue is being transcribed and it is desirable to retain formatting of the reference (e.g. superscript or italic text) then the <hi> element may be used. If encoding multiple semantic divisions in a single location reference then a nested <locus> may be used to separate or annotate these.

Finally, the content of the <locus> element may be omitted if a formatting application can construct it automatically from the values of the *from* and *to* attributes:

```
<msItem  n="1">
  <locusGrp>
    <locus  from="1r"  to="12v"/>
    <locus  from="18"  to="24r"/>
  </locusGrp>
  <title>Apocalypsis beati Ioannis Apostoli</title>
</msItem>
```

#### 11.3.5.2 Linking a Location to a Transcription or an Image

The <locus> attribute can also be used to associate a location within a manuscript with facsimile images of that location, using the *facs* attribute, or with a transcription of the text occurring at that location. The former association is effected by means of the *facs* attribute; the latter by means of the *target* attribute.

The *facs* is available only when the `transcr` module described in chapter *12. Representation of Primary Sources* is included in a schema. It associates a <locus> element with one or more digitized images, as in the following example:

```
<msItem>
  <locus  facs="images/08v.jpg images/09r.jpg images/09v.jpg images/10r.jpg
images/10v.jpg">fols. 8v-10v</locus>
  <title>Birds Praise of Love</title>
  <bibl>
    <title>IMEV</title>
    <biblScope>1506</biblScope>
  </bibl>
</msItem>
```

Here, the *facs* attribute uses a URI reference to point directly to images of the relevant pages. This method may be found cumbersome when many images are to be associated with a single location. It is of most use when specific pages are referenced within a description, as in the following example:

```
<decoDesc>
  <p>Several of the miniatures in this section have been damaged and overpainted
     at a later date (e.g. the figure of Christ on <locus  facs="http://www.example.com/images.fr#
33r</locus>; the face of the
     Shepherdess on <locus  facs="http://www.example.com/images.fr#F59V">fol.
        59v</locus>, etc.).</p>
</decoDesc>
```

For further discussion of the *facs* attribute, see section *12.1. Digital Facsimiles*.

Where a transcription of the relevant pages is available, this may be associated with the <locus> element using its *target* attribute, as in the following example:

```
<msItem  n="1">
  <locus  target="#f1r #f1v #f2r">ff. 1r-2r</locus>
  <author>Ben Jonson</author>
  <title>Ode to himself</title>
  <rubric  rend="italics">
    <lb/> An Ode
  <lb/> to him selfe.</rubric>
  <incipit>Com leaue the loathed stage</incipit>
  <explicit>And see his chariot triumph ore his wayne.</explicit>
  <bibl>
    <name>Beal</name>, <title>Index 1450-1625</title>, JnB 380</bibl>
</msItem>
<!-- within transcription ... -->
<pb  xml:id="f1r"/>
<!-- ... -->
<pb  xml:id="f1v"/>
<!-- ... -->
<pb  xml:id="f2r"/>
<!-- ... -->
```

When (as in this example) a sequence of elements is to be supplied as target value, it may be given explicitly as above, or using the xPointer range() syntax defined at *17.2.4.6. range()*. Note however that support for this pointer mechanism is not widespread in current XML processing systems.

The *target* attribute should only be used to point to elements that contain or indicate a transcription of the locus being described. To associate a <locus> element with a page image or other comparable representation, the global *facs* attribute should be used instead.

#### 11.3.5.3 Using Multiple Location Schemes

Where a manuscript contains more than one foliation, the *scheme* attribute may be used to distinguish them. For example, MS 65 Corpus Christi College, Cambridge contains two fly leaves bearing music. These leaves have modern foliation 135 and 136 respectively, but are also marked with an older foliation. This may be preserved in an encoding such as the following:

```
<locus  scheme="#original">XCIII</locus>
<locus  scheme="#modern">135</locus>
```

Here the *scheme* attribute points to a <foliation> element providing more details about the scheme used, as further discussed in *11.7.1.4. Foliation* below.

Where discontinuous sequences are identified within two different foliations, the *scheme* attribute should be supplied on the <locusGrp> element in preference, as in the following:

```
<locusGrp scheme="#original">
  <locus>XCIII</locus>
  <locus>CC-CCII</locus>
</locusGrp>
<locusGrp scheme="#modern">
  <locus>135</locus>
  <locus>197-204</locus>
</locusGrp>
```

### 11.3.6 Names of Persons, Places, and Organizations

The standard TEI element <name> may be used to identify names of any kind occurring within a description:

`<name>` (name, proper noun) contains a proper noun or noun phrase.

As further discussed in *3.6.1. Referring Strings*, this element is a member of the class `att.canonical`, from which it inherits the following attributes:

`att.canonical` provides attributes that can be used to associate a representation such as a name or title with canonical information about the object being named or referenced.

- `@key` provides an externally-defined means of identifying the entity (or entities) being named, using a coded value of some kind.
- `@ref` (reference) provides an explicit means of locating a full definition or identity for the entity being named by means of one or more URIs.

Here are some examples of the use of the <name> element:

```
<name type="person">Thomas Hoccleve</name>
<name type="place">Villingaholt</name>
<name type="org">Vetus Latina Institut</name>
<name type="person" ref="#HOC001">Occleve</name>
```

Note that the <name> element is defined as providing information about a *name*, not the person, place, or organization to which that name refers. In the last example above, the *ref* attribute is used to associate the name with a more detailed description of the person named. This is provided by means of the <person> element, which becomes available when the `namesdates` module described in chapter *14. Names, Dates, People, and Places* is included in a schema. An element such as the following might then be used to provide detailed information about the person indicated by the name:

```
<person xml:id="HOC001">
  <persName>
    <surname>Hoccleve</surname>
    <forename>Thomas</forename>
  </persName>
  <birth notBefore="1368"/>
  <occupation>poet</occupation>
<!-- other personal data -->
</person>
```

Note that an instance of the <person> element must be provided for each distinct *ref* value specified. For example, in the case above, the value HOC001 must be found as the *xml:id* attribute of some <person> element; the same value will be used as the *ref* attribute of every reference to Hoccleve in the document (however spelled), but there will only be one <person> element with this identifier.

Alternatively, the *key* attribute may be used to supply a unique identifying code for the person referenced by the name independently of both the existence of a <person> element and the use of the standard URI reference mechanism. If, for example, a project maintains as its authority file some non-digital resource, or uses a database

which cannot readily be integrated with other digital resources for this purpose, the unique codes used by such 'offline' resources may be used as values for the *key* attribute. Although such practices clearly reduce the interchangeability of the resulting encoded texts, they may be judged more convenient or practical in certain situations. As explained in *3.6.1. Referring Strings*, interchange is improved by use of tag URIs in *ref* instead of *key*.

All the <person> elements referenced by a particular document set should be collected together within a <listPerson> element, located in a <standOff> element. This functions as a kind of prosopography for all the people referenced by the set of manuscripts being described, in much the same way as a <listBibl> element may be used to hold bibliographic information for all the works referenced.

When the `namesdates` module described in chapter *14. Names, Dates, People, and Places* is included in a schema, similar mechanisms are used to maintain and reference canonical lists of places or organizations, as further discussed in sections *14.2.3. Place Names* and *14.2.2. Organizational Names* respectively.

### 11.3.7 Catchwords, Signatures, Secundo Folio

The <catchwords> element is used to describe one method by which correct ordering of the quires of a codex is ensured. Typically, this takes the form of a word or phrase written in the lower margin of the last leaf verso of a gathering, which provides a preview of the first recto leaf of the successive gathering. This may be a simple phrase such as the following:

```
<catchwords>Quires signed on the
last leaf verso in roman numerals.</catchwords>
```

Alternatively, it may contain more details:

```
<catchwords>Vertical catchwords in the hand of the scribe placed along the inner
bounding line, reading from top to bottom.</catchwords>
```

The 'Signatures' element is used, in a similar way, to describe a similar system in which quires or leaves are marked progressively in order to facilitate arrangement during binding. For example:

```
<signatures>At the bottom of the
first four leaves of quires 1-14 are the remains of a series of quire signatures
a-o plus roman figures in a cursive hand of the fourteenth century.</signatures>
```

The <signatures> element can be used for either leaf signatures, or a combination of quire and leaf signatures, whether the marking is alphabetic, alphanumeric, or some ad hoc system, as in the following more complex example:

```
<signatures>Quire and leaf
signatures in letters, [b]-v, and roman numerals; those in quires 10 (1) and 17
(s) in red ink and different from others; every third quire also signed with red
crayon in arabic numerals in the centre lower margin of the first leaf recto:
"2" for quire 4 (f. 19), "3" for quire 7 (f. 43); "4", barely visible, for quire
10 (f. 65), "5", in a later hand, for quire 13 (f. 89), "6", in a later hand,
for quire 16 (f. 113).</signatures>
```

The <secFol> element (for 'secundo folio') is used to record an identifying phrase (also called *dictio probatoria*) taken from a specific known point in a codex (for example the first few words on the second leaf). Since these words will differ from one copy of a text to another, the practice originated in the middle ages of using them when cataloguing a manuscript in order to distinguish individual copies of a work in a way which its opening words could not.

```
<secFol>(ando-)ssene in una villa</secFol>
```

### 11.3.8 Heraldry

Descriptions of heraldic arms, supporters, devices, and mottos may appear at various points in the description of a manuscript, usually in the context of ownership information, binding descriptions, or detailed accounts of illustrations. A full description may also contain a detailed account of the heraldic components of a manuscript independently considered. Frequently, however, heraldic descriptions will be cited as short phrases within other parts of the record. The phrase level element <heraldry> is provided to allow such phrases to be marked for further analysis, as in the following examples:

```
<p>Ownership stamp (xvii cent.) on i recto with the arms <heraldry>A bull
    passant within a bordure bezanty, in chief a crescent for difference</heraldry>
[Cole], crest, and the legend <quote>Cole Deum</quote>.</p>
<!-- ... -->
<p>A c. 8r fregio su due lati, <heraldry>stemma e imprese medicee</heraldry>
racchiudono l'inizio dell'epistolario di Paolino.</p>
```

## 11.4 The Manuscript Identifier

The <msIdentifier> element is intended to provide an unambiguous means of uniquely identifying a particular manuscript. This may be done in a structured way, by providing information about the holding institution and the call number, shelfmark, or other identifier used to indicate its location within that institution. Alternatively, or in addition, a manuscript may be identified simply by a commonly used name.

`<msIdentifier>` (manuscript identifier) contains the information required to identify the manuscript or similar object being described.

A manuscript's actual physical location may occasionally be different from its place of ownership; at Cambridge University, for example, manuscripts owned by various colleges are kept in the central University Library. Normally, it is the ownership of the manuscript which should be specified in the manuscript identifier, while additional or more precise information on the physical location of the manuscript can be given within the <adminInfo> element, discussed in section *11.9.1. Administrative Information* below.

The following elements are available within <msIdentifier> to identify the holding institution:

`<country>` (country) contains the name of a geo-political unit, such as a nation, country, colony, or commonwealth, larger than or administratively superior to a region and smaller than a bloc.

`<region>` (region) contains the name of an administrative unit such as a state, province, or county, larger than a settlement, but smaller than a country.

`<settlement>` (settlement) contains the name of a settlement such as a city, town, or village identified as a single geo-political or administrative unit.

`<institution>` (institution) contains the name of an organization such as a university or library, with which a manuscript or other object is identified, generally its holding institution.

`<repository>` (repository) contains the name of a repository within which manuscripts or other objects are stored, possibly forming part of an institution.

These elements are all structurally equivalent to the standard TEI <name> element with an appropriate value for its *type* attribute; however the use of this 'syntactic sugar' enables the model for <msIdentifier> to be constrained rather more tightly than would otherwise be possible. Specifically, only one of each of the elements listed above may appear within the <msIdentifier> and they must, if present, appear in the order given.

Like <name>, these elements are all also members of the attribute class `att.canonical`, and thus can use the attributes *key* or *ref* to reference a single standardized source of information about the entity named.

The following elements are used within <msIdentifier> to provide different ways of identifying the manuscript within its holding institution:

`<collection>` (collection) contains the name of a collection of manuscripts or other objects, not necessarily located within a single repository.

`<idno>` (identifier) supplies any form of identifier used to identify some object, such as a bibliographic item, a person, a title, an organization, etc. in a standardized way.

`<altIdentifier>` (alternative identifier) contains an alternative or former structured identifier used for a manuscript or other object, such as a former catalogue number.

`<msName>` (manuscript name) contains a proper noun or noun phrase used for a manuscript, or other object, as opposed to a formal identification number or classmark.

Major manuscript repositories will usually have a preferred form of citation for manuscript shelfmarks, including rules about punctuation, spacing, abbreviation, etc., which should be adhered to. Where such a format also contains information which might additionally be supplied as a distinct subcomponent of the <msIdentifier>, for example a collection name, a decision must be taken as to whether to use the more specific element, or to include such information within the <idno> element. For example, the manuscript formally identified as 'El 26 C 9' forms a part of the Ellesmere ('El') collection. Either of the following encodings is therefore feasible:

```
<msIdentifier>
  <country>USA</country>
  <region>California</region>
  <settlement>San Marino</settlement>
  <repository>Huntington Library</repository>
  <collection>El</collection>
  <idno>26 C 9</idno>
  <msName>The Ellesmere Chaucer</msName>
</msIdentifier>
```

```
<msIdentifier>
  <country>USA</country>
  <region>California</region>
  <settlement>San Marino</settlement>
  <repository>Huntington Library</repository>
  <idno>El 26 C 9</idno>
  <msName>The Ellesmere Chaucer</msName>
</msIdentifier>
```

In the former example, the preferred form of the identifier can be retrieved by prefixing the content of the <idno> element with that of the <collection> element, while in the latter it is given explicitly. The advantage of the former is that it simplifies accurate retrieval of all manuscripts from a given collection; the disadvantage is that encoded abbreviations of this kind may not be as immediately comprehensible. Care should be taken to avoid redundancy: for example

```
<collection>El</collection>
<idno>El 26 C 9</idno>
```

would clearly be inappropriate. Equally clearly,

```
<collection>Ellesmere</collection>
<idno>El 26 C 9</idno>
```

might be considered helpful in some circumstances (if, for example, some of the items in the Ellesmere collection had shelfmarks which did not begin 'El').

In some cases the shelfmark may contain no information about the collection; in other cases, the item may be regarded as belonging to more than one collection. The <collection> element may be added, and repeated as often as necessary to cater for such situations:

```
<msIdentifier>
  <country>Hungary</country>
  <settlement>Budapest</settlement>
  <repository  xml:lang="fr"> Bibliothèque de l'Académie des Sciences de Hongrie
</repository>
  <collection>Oriental Collection</collection>
  <collection>Sandor Kégl Bequest</collection>
  <idno>MS 1265</idno>
</msIdentifier>
```

```
<msIdentifier>
  <country>USA</country>
  <region>New Jersey</region>
  <settlement>Princeton</settlement>
  <repository>Princeton University Library</repository>
  <collection>Scheide Library</collection>
  <idno>MS 71</idno>
  <msName>Blickling Homiliary</msName>
</msIdentifier>
```

Note in the latter case the use of the <msName> element to provide a common name other than the shelfmark by which a manuscript is known. Where a manuscript has several such names, more than one of these elements may be used, as in the following example:

```
<msIdentifier>
  <country>Danmark</country>
  <settlement>København</settlement>
  <repository>Det Arnamagnæanske Institut</repository>
  <idno>AM 45 fol.</idno>
  <msName  xml:lang="la">Codex Frisianus</msName>
  <msName  xml:lang="is">Fríssbók</msName>
</msIdentifier>
```

Here the globally available *xml:lang* attribute has been used to specify the language of the alternative names.

In very rare cases a repository may have only one manuscript (or only one of any significance), which will have no shelfmark as such but will be known by a particular name or names. In such circumstances, the <idno> element may be omitted, and the manuscript identified by the name or names used for it, using one or more <msName> elements, as in the following example:

```
<msIdentifier>
  <settlement>Rossano</settlement>
  <repository  xml:lang="it">Biblioteca arcivescovile</repository>
  <msName  xml:lang="la">Codex Rossanensis</msName>
  <msName  xml:lang="la">Codex purpureus</msName>
  <msName  xml:lang="en">The Rossano Gospels</msName>
</msIdentifier>
```

If a manuscript name contains a name or referencing string that it is useful to annotate (e.g. by referring to an authority list) then <name> or <rs> may be used for this purpose.

Where manuscripts have moved from one institution to another, or even within the same institution, they may have identifiers additional to the ones currently used, such as former shelfmarks, which are sometimes retained even

after they have been officially superseded. In such cases it may be useful to supply an alternative identifier, with a detailed structure similar to that of the <msIdentifier> itself. The following example shows a manuscript which had shelfmark `II-M-5` in the collection of the Duque de Osuna, but which now has the shelfmark `MS 10237` in the National Library in Madrid:

```
<msIdentifier>
  <settlement>Madrid</settlement>
  <repository>Biblioteca Nacional</repository>
  <idno>MS 10237</idno>
  <altIdentifier>
    <region>Andalucia</region>
    <settlement>Osuna</settlement>
    <repository>Duque de Osuna</repository>
    <idno>II-M-5</idno>
  </altIdentifier>
</msIdentifier>
```

Normally, such information would be dealt with under <history>, except in cases where a manuscript is likely still to be referred to or known by its former identifier. For example, an institution may have changed its call number system but still wish to retain a record of the earlier number, perhaps because the manuscript concerned is frequently cited in print under its previous number:

```
<msIdentifier>
  <settlement>Berkeley</settlement>
  <institution>University of California</institution>
  <repository>Bancroft Library</repository>
  <idno>UCB 16</idno>
  <altIdentifier>
    <idno>2MS BS1145 I8</idno>
  </altIdentifier>
</msIdentifier>
```

Where (as in this example) no repository is specified for the <altIdentifier>, it is assumed to be the same as that of the parent <msIdentifier>. Where the holding institution has only one preferred form of citation but wishes to retain the other for internal administrative purposes, the secondary could be given within <altIdentifier> with an appropriate value on the *type* attribute:

```
<msIdentifier>
  <settlement>Oxford</settlement>
  <repository>Bodleian Library</repository>
  <idno>MS. Bodley 406</idno>
  <altIdentifier type="SC">
    <idno>2297</idno>
  </altIdentifier>
</msIdentifier>
```

It might, however, be preferable to include such information within the <adminInfo> element discussed in section *11.9.1. Administrative Information* below.

Cases of such changed or alternative identifiers should be clearly distinguished from cases of 'fragmented' (*11.11. Manuscript Fragments*) manuscripts, that is to say manuscripts which although physically disjoint are nevertheless generally treated as single units.

As mentioned above, the smallest possible description is one that contains only the element <msIdentifier>; good practice in all but exceptional circumstances requires the presence within it of the three sub-elements <settlement>, <repository>, and <idno>, since they provide what is, by common consent, the minimum amount of information necessary to identify a manuscript.

## 11.5 The Manuscript Heading

Historically, the briefest possible meaningful description of a manuscript consists of no more than a title, e.g. *Polychronicon*. This will often have been enough to identify a manuscript in a small collection because the identity of the author is implicit. Where a title does not imply the author, and is thus insufficient to identify the main text of a manuscript, the author should be stated explicitly (e.g. *Augustinus, Sermones* or *Cicero, Letters*). Many inventories of manuscripts consist of no more than an author and title, with some form of copy-specific identifier, such as a shelfmark or 'secundo folio' reference (e.g. *Arch. B. 3. 2: Evangelium Matthei cum glossa*, *126. Isidori Originum libri octo*, *Biblia Hieronimi, 2o fo. opus est*); information on date and place of writing will sometimes also be included. The standard TEI element <head> element can be used to provide a brief description of this kind.

`<head>` (heading) contains any type of heading, for example the title of a section, or the heading of a list, glossary, manuscript description, etc.

In this way the cataloguer or scholar can supply in one place a minimum of essential information, such as might be displayed or printed as the heading of a full description. For example:

```
<head>Marsilius de Inghen, Abbreviata phisicorum Aristotelis; Italy,
1463.</head>
```

Any phrase-level elements, such as <title>, <name>, <date>, or the specialized elements <origPlace> and <origDate>, can also be used within a <head> element, but it should be remembered that the <head> element is intended principally to contain a heading. More structured information concerning the contents, physical form, or history of the manuscript should be given within the specialized elements described below, <msContents>, <physDesc>, <history>, etc. However, in simple cases, the <p> element may also be used to supply an unstructured collection of such information, as in the example given above (*11.2. The Manuscript Description Element*).

## 11.6 Intellectual Content

The <msContents> element is used to describe the intellectual content of a manuscript or manuscript part. It comprises *either* a series of informal prose paragraphs *or* a series of <msItem> or <msItemStruct> elements, each of which provides a more detailed description of a single item contained within the manuscript. These may be prefaced, if desired, by a <summary> element, which is especially useful where one wishes to provide an overview of a manuscript's contents and describe only some of the items in detail.

`<msContents>` (manuscript contents) describes the intellectual content of a manuscript, manuscript part, or other object either as a series of paragraphs or as a series of structured manuscript items.

`<msItem>` (manuscript item) describes an individual work or item within the intellectual content of a manuscript, manuscript part, or other object.

`<msItemStruct>` (structured manuscript item) contains a structured description for an individual work or item within the intellectual content of a manuscript, manuscript part, or other object.

`<summary>` contains an overview of the available information concerning some aspect of an item or object (for example, its intellectual content, history, layout, typography etc.) as a complement or alternative to the more detailed information carried by more specific elements.

In the simplest case, only a brief description may be provided, as in the following examples:

```
<msContents>
  <p>A collection of Lollard sermons</p>
</msContents>
```

```
<msContents>
  <p>Atlas of the world from Western Europe and Africa to Indochina, containing 27
     maps and 26 tables</p>
</msContents>
```

```
<msContents>
  <p>Biblia sacra: Antiguo y Nuevo Testamento, con prefacios, prólogos y
     argumentos de san Jerónimo y de otros. Interpretaciones de los nombres
     hebreos.</p>
</msContents>
```

This description may of course be expanded to include any of the TEI elements generally available within a <p> element, such as <title>, <bibl>, or <list>. More usually, however, each individual work within a manuscript will be given its own description, using the <msItem> or <msItemStruct> element described in the next section, as in the following example:

```
<msContents>
  <msItem  n="1">
    <locus>fols. 5r -7v</locus>
    <title>An ABC</title>
    <bibl>
      <title>IMEV</title>
      <biblScope>239</biblScope>
    </bibl>
  </msItem>
  <msItem  n="2">
    <locus>fols. 7v -8v</locus>
    <title  xml:lang="fr">Lenvoy de Chaucer a Scogan</title>
    <bibl>
      <title>IMEV</title>
      <biblScope>3747</biblScope>
    </bibl>
  </msItem>
  <msItem  n="3">
    <locus>fol. 8v</locus>
    <title>Truth</title>
    <bibl>
      <title>IMEV</title>
      <biblScope>809</biblScope>
    </bibl>
  </msItem>
  <msItem  n="4">
    <locus>fols. 8v-10v</locus>
    <title>Birds Praise of Love</title>
    <bibl>
      <title>IMEV</title>
      <biblScope>1506</biblScope>
    </bibl>
  </msItem>
  <msItem  n="5">
    <locus>fols. 10v -11v</locus>
    <title  xml:lang="la">De amico ad amicam</title>
    <title  xml:lang="la">Responcio</title>
    <bibl>
      <title>IMEV</title>
      <biblScope>16 & 19</biblScope>
    </bibl>
  </msItem>
  <msItem  n="6">
    <locus>fols. 14r-126v</locus>
    <title>Troilus and Criseyde</title>
```

```
    <note>Bk. 1:71-Bk. 5:1701, with additional losses due to mutilation
        throughout</note>
  </msItem>
</msContents>
```

The <summary> element may be used in conjunction with one or more <msItem> elements if it is desired to provide both a general summary of the contents of a manuscript and more detail about some or all of the individual items within it. It may not however be used within an individual <msItem> element.

```
<msContents>
  <summary>A collection of Lollard sermons</summary>
  <msItem  n="1">
    <locus>fol. 4r-8r</locus>
    <title>3rd Sunday Before Lent</title>
  </msItem>
  <msItem  n="2">
    <locus>fol. 9r-16v</locus>
    <title>Sexagesima</title>
  </msItem>
</msContents>
```

### 11.6.1 The <msItem> and <msItemStruct> Elements

Each discrete item in a manuscript or manuscript part can be described within a distinct <msItem> or <msItemStruct> element, and may be classified using the *class* attribute.

These are the possible component elements of <msItem> and <msItemStruct>.

`<author>` (author) in a bibliographic reference, contains the name(s) of an author, personal or corporate, of a work; for example in the same form as that provided by a recognized bibliographic name authority.

`<respStmt>` (statement of responsibility) supplies a statement of responsibility for the intellectual content of a text, edition, recording, or series, where the specialized elements for authors, editors, etc. do not suffice or do not apply. May also be used to encode information about individuals or organizations which have played a role in the production or distribution of a bibliographic work.

`<title>` (title) contains a title for any kind of work.

`@type` classifies the title according to some convenient typology.

`<rubric>` (rubric) contains the text of any *rubric* or heading attached to a particular manuscript item, that is, a string of words through which a manuscript or other object signals the beginning of a text division, often with an assertion as to its author and title, which is in some way set off from the text itself, typically in red ink, or by use of different size or type of script, or some other such visual device.

`<incipit>` contains the *incipit* of a manuscript or similar object item, that is the opening words of the text proper, exclusive of any *rubric* which might precede it, of sufficient length to identify the work uniquely; such incipits were, in former times, frequently used a means of reference to a work, in place of a title.

`<quote>` (quotation) contains a phrase or passage attributed by the narrator or author to some agency external to the text.

`<explicit>` (explicit) contains the *explicit* of a item, that is, the closing words of the text proper, exclusive of any rubric or colophon which might follow it.

`<finalRubric>` (final rubric) contains the string of words that denotes the end of a text division, often with an assertion as to its author and title, usually set off from the text itself by red ink, by a different size or type of script, or by some other such visual device.

`<colophon>` (colophon) contains the *colophon* of an item: that is, a statement providing information regarding the date, place, agency, or reason for production of the manuscript or other object.

`<decoNote>` (note on decoration) contains a note describing either a decorative component of a manuscript or other object, or a fairly homogenous class of such components.

`<listBibl>` (citation list) contains a list of bibliographic citations of any kind.

`<bibl>` (bibliographic citation) contains a loosely-structured bibliographic citation of which the sub-components may or may not be explicitly tagged.

`<filiation>` (filiation) contains information concerning the manuscript or other object's *filiation*, i.e. its relationship to other surviving manuscripts or other objects of the same text or contents, its *protographs*, *antigraphs* and *apographs*.

`<note>` (note) contains a note or annotation.

`<textLang>` (text language) describes the languages and writing systems identified within the bibliographic work being described, rather than its description.

In addition, an <msItemStruct> may contain nested <msItemStruct> elements, just as an <msItem> may contain nested <msItem> elements.

The main difference between <msItem> and <msItemStruct> is that in the former, the order and number of child elements is not constrained; any element, in other words, may be given in any order, and repeated as often as is judged necessary. In the latter, however, the sub-elements, if used, must be given in the order specified above and only some of them may be repeated; specifically, <rubric>, <finalRubric>. <incipit>, <textLang> and <explicit> can appear only once.

While neither <msItem> nor <msItemStruct> may contain untagged running text, both permit an unstructured description to be provided in the form of one or more paragraphs of text. They differ in this respect also: if paragraphs are supplied as the content of an <msItem>, then none of the other component elements listed above is permitted; in the <msItemStruct> case, however, paragraphs may appear anywhere as an alternative to any of the component elements listed above.

As noted above, both <msItem> and <msItemStruct> elements may also nest, where a number of separate items in a manuscript are grouped under a single title or rubric, as is the case, for example, with a work like *The Canterbury Tales*.

The elements <msContents>, <msItem>, <msItemStruct>, <incipit>, and <explicit> are all members of the class `att.msExcerpt` from which they inherit the *defective* attribute.

`att.msExcerpt` (manuscript excerpt) provides attributes used to describe excerpts from a manuscript placed in a description thereof.

- `@defective` indicates whether the passage being quoted is defective, i.e. incomplete through loss or damage.

This attribute can be used for example with collections of fragments, where each fragment is given as a separate <msItem> and the first and last words of each fragment are transcribed as defective incipits and explicits, as in the following example, a manuscript containing four fragments of a single work:

```
<msContents>
  <msItem defective="true">
    <locus from="1r" to="9v">1r-9v</locus>
    <title>Knýtlinga saga</title>
    <msItem n="1.1">
      <locus from="1r:1" to="2v:30">1r:1-2v:30</locus>
      <incipit defective="true">dan<ex>n</ex>a a engl<ex>an</ex>di</incipit>
      <explicit defective="true">en meðan <expan>haraldr</expan> hein hafði
            k<ex>onung</ex>r v<am>
          <g ref="http://www.example.com/abbrevs.xml#er"/>
        </am>it
            yf<ex>ir</ex> danmork</explicit>
    </msItem>
<!-- msItems 1.2 to 1.4 -->
  </msItem>
</msContents>
```

The elements <ex>, <am>, and <expan> used in the above example are further discussed in section *12.3.1.2. Abbreviation and Expansion*; they are available only when the `transcr` module defined by that chapter is selected.

Similarly, the <g> element used in this example to represent the abbreviation mark is defined by the `gaiji` module documented in chapter *5. Characters, Glyphs, and Writing Modes*.

### 11.6.2 Authors and Titles

When used within a manuscript description, the <title> element should be used to supply a regularized form of the item's title, as distinct from any rubric quoted from the manuscript. If the item concerned has a standardized distinctive title, e.g. *Roman de la Rose*, then this should be the form given as content of the <title> element, with the value of the *type* attribute given as `uniform`. If no uniform title exists for an item, or none has been yet identified, or if one wishes to provide a general designation of the contents, then a 'supplied' title can be given, e.g. *missal*, in which case the *type* attribute on the <title> should be given the value `supplied`.

Similarly, if used within a manuscript description, the <author> element should always contain the normalized form of an author's name, irrespective of how (or whether) this form of the name is cited in the manuscript. If it is desired to retain the form of the author's name as given in the manuscript, this may be tagged as a distinct <name> element, within the text at the point where it occurs.

Note that the *key* attribute can also be used, as on names in general, to specify the identifier of a <person> element carrying full details of the person concerned (see further *11.3.6. Names of Persons, Places, and Organizations*).

The <respStmt> element can be used to supply the name and role of a person other than the author who is responsible for some aspect of the intellectual content of the manuscript:

```
<author>Diogenes Laertius</author>
<respStmt>
  <resp>in the translation of</resp>
  <name>Ambrogio Traversari</name>
</respStmt>
```

The <respStmt> element can also be used where there is a discrepancy between the author of an item as given in the manuscript and the accepted scholarly view, as in the following example:

```
<title  type="supplied">Sermons on the Epistles and the Gospels</title>
<respStmt>
  <resp>here erroneously attributed to</resp>
  <name>St. Bonaventura</name>
</respStmt>
```

Note that such attributions of authorship, both correct and incorrect, are frequently found in the rubric or final rubric (and occasionally also elsewhere in the text), and can therefore be transcribed and included in the description, if desired, using the <rubric>, <finalRubric>, or <quote> elements, as appropriate.

### 11.6.3 Rubrics, Incipits, Explicits, and Other Quotations from the Text

It is customary in a manuscript description to record the opening and closing words of a text as well as any headings or colophons it might have, and the specialized elements <rubric>, <incipit>, <explicit>, <finalRubric>, and <colophon> are available within <msItem> for doing so, along with the more general <quote>, for recording other bits of the text not covered by these elements. Each of these elements has the same substructure, containing a mixture of phrase-level elements and plain text. A <locus> element can be included within each, in order to specify the location of the component, as in the following example:

```
<msContents>
  <msItem>
    <locus>f. 1-223</locus>
    <author>Radulphus Flaviacensis</author>
    <title>Expositio super Leviticum </title>
```

```
    <incipit>
      <locus>f. 1r</locus> Forte Hervei monachi</incipit>
    <explicit>
      <locus>f. 223v</locus> Benedictio salis et aquae</explicit>
  </msItem>
</msContents>
```

In the following example, standard TEI elements for the transcription of primary sources have been used to mark the expansion of abbreviations and other features present in the original:

```
<msItem  defective="true">
  <locus>ff. 1r-24v</locus>
  <title  type="uniform">Ágrip af Noregs konunga sǫgum</title>
  <incipit  defective="true">
    <lb/>regi oc h<ex>ann</ex> seti ho<gap  reason="illegible"  quantity="7"
      unit="mm"/>
    <lb/>sc heim se<ex>m</ex> þio</incipit>
  <explicit  defective="true">h<ex>on</ex> hev<ex>er</ex>
    <ex>oc</ex> þa buit hesta .ij.
  <lb/>annan viþ fé en h<ex>on</ex>o<ex>m</ex> annan til
     reiþ<ex>ar</ex>
  </explicit>
</msItem>
```

Note here also the use of the *defective* attribute on <incipit> and <explicit> to indicate that the text begins and ends defectively.

The *xml:lang* attribute for <colophon>, <explicit>, <incipit>, <quote>, and <rubric> may always be used to identify the language of the text quoted, if this is different from the default language specified by the *mainLang* attribute on <textLang>.

### 11.6.4 Filiation

The <filiation> element can be used to provide information on the relationship between the manuscript and other surviving manuscripts of the same text, either specifically or in a general way, as in the following example:

```
<msItem>
  <locus>118rb</locus>
  <incipit>Ecce morior cum nichil horum ... <ref>[Dn 13, 43]</ref>. Verba ista
     dixit Susanna de illis</incipit>
  <explicit>ut bonum comune conservatur.</explicit>
  <bibl>Schneyer 3, 436 (Johannes Contractus OFM)</bibl>
  <filiation>weitere Überl. Uppsala C 181, 35r.</filiation>
</msItem>
```

### 11.6.5 Text Classification

One or more text classification or text-type codes may be specified, either for the whole of the <msContents> element, or for one or more of its constituent <msItem> elements, using the *class* attribute as specified above:

```
<msContents>
  <msItem  n="1"  defective="false"
    class="#law">
    <locus  from="1v"  to="71v">1v-71v</locus>
    <title  type="uniform">Jónsbók</title>
```

```
    <incipit>Magnus m<ex>ed</ex> guds miskun Noregs k<ex>onungu</ex>r</incipit>
    <explicit>en<ex>n</ex> u<ex>ir</ex>da þo t<ex>il</ex> fullra aura</explicit>
  </msItem>
</msContents>
```

The value used for the *class* attribute in this example points to a <category> element with the identifier `law`, which defines the classification concerned. Such <category> elements will typically appear within a <taxonomy> element, within the <classDecl> element of the TEI header (*2.3.7. The Classification Declaration*) as in the following example:

```
<classDecl>
  <taxonomy>
<!-- -->
    <category  xml:id="law">
      <catDesc>Legislation</catDesc>
    </category>
    <category  xml:id="war">
      <catDesc>Military topics</catDesc>
    </category>
<!-- -->
  </taxonomy>
</classDecl>
```

More than one classification may apply to a single item. Another text, concerned with legislation about military topics might thus be specified as follows:

```
<msItem  class="#law #war">
  <p>A treatise on Clausewitz</p>
<!-- details of the item here -->
</msItem>
```

### 11.6.6 Languages and Writing Systems

The <textLang> element should be used to provide information about the languages used within a manuscript item. It may take the form of a simple note, as in the following example:

```
<textLang>Old Church Slavonic, written in Cyrillic script.</textLang>
```

Where, for validation and indexing purposes, it is thought convenient to add keywords identifying the particular languages used, the *mainLang* attribute may be used. This attribute takes the same range of values as the global *xml:lang* attribute, on which see further *vi.1 Language Identification.* In the following example a manuscript written chiefly in Old Church Slavonic is described:

```
<textLang  mainLang="cu">Old Church Slavonic</textLang>
```

A manuscript item will sometimes contain material in more than one language. The *mainLang* attribute should be used only for the chief language. Other languages used may be specified using the *otherLangs* attribute as in the following example:

```
<textLang  mainLang="cu"  otherLangs="ru grc">Mostly Old Church Slavonic, with
some Russian and Greek material</textLang>
```

Since Old Church Slavonic may be written in either Cyrillic or Glagolitic scripts, and even occasionally in both within the same manuscript, it might be preferable to use a more explicit identifier:

```
<textLang mainLang="cu-Cyrs">Old Church Slavonic in Cyrillic script</textLang>
```

The form and scope of language identifiers recommended by these Guidelines is based on the IANA standard described at *vi.1 Language Identification* and should be followed throughout. Where additional detail is needed correctly to describe a language, or to discuss its deployment in a given text, this should be done using the <langUsage> element in the TEI header, within which individual <language> elements document the languages used: see *2.4.2. Language Usage*.

Note that the <language> element defines a particular combination of human language and writing system. Only one <language> element may be supplied for each such combination. Standard TEI practice also allows this element to be referenced by any element using the global *xml:lang* attribute in order to specify the language applicable to the content of that element. For example, assuming that <language> elements have been defined with the identifiers `fr` (for French), `la` (for Latin), and `de` (for German), a manuscript description written in French which specifies that a particular manuscript contains predominantly German but also some Latin material, might have a <textLang> element like the following:

```
<textLang xml:lang="fr" mainLang="de"
 otherLangs="la">allemand et
latin</textLang>
```

## 11.7 Physical Description

Under the general heading 'physical description' we subsume a large number of different aspects generally regarded as useful in the description of a given manuscript. These include:

- aspects of the form, support, extent, and quire structure of the manuscript object and of the way in which the text is laid out on the page (*11.7.1. Object Description*);
- the styles of writing, such as the way it is laid out on the page, the styles of writing, decorative features, any musical notation employed and any annotations or marginalia (*11.7.2. Writing, Decoration, and Other Notations*);
- and discussion of its binding, seals, and any accompanying material (*11.7.3. Bindings, Seals, and Additional Material*).

Most manuscript descriptions touch on several of these categories of information though few include them all, and not all distinguish them as clearly as we propose here. In particular, it is often the case that an existing description will include information for which we propose distinct elements within a single paragraph, or even sentence. The encoder must then decide whether to rewrite the description using the structure proposed here, or to retain the existing prose, marked up simply as a series of <p> elements, directly within the <physDesc> element.

The <physDesc> element may thus be used in either of two distinct ways. It may contain a series of paragraphs addressing topics listed above and similar ones. Alternatively, it may act as a container for any choice of the more specialized elements described in the remainder of this section, each of which itself contains a series of paragraphs, and may also have more specific attributes.

In general, it is not recommended to combine unstructured prose description with usage of the more specialized elements, as such an approach complicates processing, and may lead to inconsistency within a single manuscript description. A single <physDesc> element will normally contain either a series of `model.pLike` elements, or a sequence of specialized elements from the `model.physDescPart` class. There are however circumstances in which this is not feasible, for example:

- the description already exists in a prose form where some of the specialized topics are treated together in paragraphs of prose, but others are treated distinctly;
- although all parts of the description are clearly distinguished, some of them cannot be mapped to a pre-existing specialized element.

In such situations, both specialized and generic (`model.pLike`) elements may be combined in a single <physDesc>. Note however that all generic elements given must precede the first specialized element in the description. Thus the following is valid:

```
<physDesc>
  <p>Generic descriptive prose...</p>
<!-- other generic elements here -->
  <objectDesc  form="codex">
<!-- ... -->
  </objectDesc>
<!-- other specific elements here -->
</physDesc>
```

but neither of the following is valid:

```
<physDesc>
<objectDesc form="codex">
<!-- ... -->
</objectDesc>
<p>Generic descriptive prose...</p>
</physDesc>
```

```
<physDesc>
<objectDesc form="codex">
  <!-- ... -->
</objectDesc>
<p>Generic descriptive prose...</p>
<!-- other specific elements here -->
</physDesc>
```

The order in which specific elements may appear is also constrained by the content model; again this is for simplicity of processing. They may of course be processed or displayed in any desired order, but for ease of validation, they must be given in the order specified below.

### 11.7.1 Object Description

The <objectDesc> element is used to group together those parts of the physical description which relate specifically to the text-bearing object, its format, constitution, layout, etc. The <objectDesc> element is used for grouping elements relating to the physicality of a text-bearing object as part of a manuscript description. If a full description of an object (text-bearing or not) is desired, the more general <object> element may be preferred.

The *form* attribute is used to indicate the specific type of writing vehicle being described, for example, as a codex, roll, tablet, etc. If used it must appear first in the sequence of specialized elements. The <objectDesc> element has two parts: a description of the *support*, i.e. the physical carrier on which the text is inscribed; and a description of the *layout*, i.e. the way text is organized on the carrier.

Taking these in turn, the description of the support is tagged using the following elements, each of which is discussed in more detail below:

`<supportDesc>` (support description) groups elements describing the physical support for the written part of a manuscript or other object.

`<support>` (support) contains a description of the materials etc. which make up the physical support for the written part of a manuscript or other object.

`<extent>` (extent) describes the approximate size of a text stored on some carrier medium or of some other object, digital or non-digital, specified in any convenient units.

<collation> (collation) contains a description of how the leaves, bifolia, or similar objects are physically arranged.

<foliation> (foliation) describes the numbering system or systems used to count the leaves or pages in a codex or similar object.

<condition> (condition) contains a description of the physical condition of the manuscript or object.

Each of these elements contains paragraphs relating to the topic concerned. Within these paragraphs, phrase-level elements (in particular those discussed above at *11.3. Phrase-level Elements*), may be used to tag specific terms of interest if so desired.

```
<objectDesc  form="codex">
  <supportDesc>
    <p>Mostly <material>paper</material>, with watermarks
    <watermark>unicorn</watermark> (<ref>Briquet 9993</ref>) and
    <watermark>ox</watermark> (close to <ref>Briquet 2785</ref>). The first and last
        leaf of each quire, with the exception of quires xvi and xviii, are constituted
        by bifolia of parchment, and all seven miniatures have been painted on inserted
        singletons of parchment.</p>
  </supportDesc>
</objectDesc>
```

This example combines information which might alternatively be more precisely tagged using the more specific elements described in the following subsections.

#### 11.7.1.1 Support

The <support> element groups together information about the physical carrier. Typically, for western manuscripts, this will entail discussion of the material (parchment, paper, or a combination of the two) written on. For paper, a discussion of any watermarks present may also be useful. If this discussion makes reference to standard catalogues of such items, these may be tagged using the standard <ref> element as in the following example:

```
<support>
  <p>
    <material>Paper</material> with watermark: <watermark>anchor in a circle with
        star on top</watermark>, <watermark>countermark B-B with trefoil</watermark>
     similar to <ref>Mošin, Anchor N 1680</ref>
    <date>1570-1585</date>.</p>
</support>
```

#### 11.7.1.2 Extent

The <extent> element, defined in the TEI header, may also be used in a manuscript description to specify the number of leaves a manuscript contains, as in the following example:

```
<extent>ii + 97 + ii</extent>
```

Information regarding the size of the leaves may be specifically marked using the phrase level <dimensions> element, as in the following example, or left as plain prose.

```
<extent>ii + 321 leaves <dimensions  unit="cm">
    <height>35</height>
    <width>27</width>
  </dimensions>
</extent>
```

Alternatively, the generic <measure> element might be used within <extent>, as in the following example:

```
<extent>
  <measure  type="composition"  unit="leaf"
    quantity="10">10 Bl.</measure>
  <measure  type="height"  quantity="37"
    unit="cm">37</measure> x <measure  type="width"  quantity="29"
    unit="cm">29</measure> cm
</extent>
```

#### 11.7.1.3 Collation

The <collation> element should be used to provide a description of a book's current and original structure, that is, the arrangement of its leaves and quires. This information may be conveyed using informal prose, or any appropriate notational convention. Although no specific notation is defined here, an appropriate element to enclose such an expression would be the <formula> element, which is provided when the `figures` module is included in a schema. Here are some examples of different ways of treating collation:

```
<collation>
  <p>
    <formula>1-3:8, 4:6, 5-13:8</formula>
  </p>
</collation>
<collation>
  <p>There are now four gatherings, the first, second and fourth originally
     consisting of eight leaves, the third of seven. A fifth gathering thought to
     have followed has left no trace. <list>
      <item>Gathering I consists of 7 leaves, a first leaf, originally conjoint with
      <locus>fol. 7</locus>, having been cut away leaving only a narrow strip along
            the gutter; the others, <locus>fols 1</locus> and <locus>6</locus>,
      <locus>2</locus> and <locus>5</locus>, and <locus>3</locus> and
      <locus>4</locus>, are bifolia.</item>
      <item>Gathering II consists of 8 leaves, 4 bifolia.</item>
      <item>Gathering III consists of 7 leaves; <locus>fols 16</locus> and
      <locus>22</locus> are conjoint, the others singletons.</item>
      <item>Gathering IV consists of 2 leaves, a bifolium.</item>
    </list>
  </p>
</collation>
<collation>
  <p>I (1, 2+9, 3+8, 4+7, 5+6, 10); II (11, 12+17, 13, 14, 15, 16, 18,
     19).</p>
</collation>
<collation>
  <p>
    <formula>1-5.8 6.6 (catchword, f. 46, does not match following text) 7-8.8
        9.10, 11.2 (through f. 82) 12-14.8 15.8(-7)</formula>
  </p>
</collation>
```

#### 11.7.1.4 Foliation

The <foliation> element may be used to indicate the scheme, medium or location of folio, page, column, or line numbers written in the manuscript, frequently including a statement about when and, if known, by whom, the numbering was done.

```
<foliation>
  <p>Neuere Foliierung, die auch das Vorsatzblatt mitgezählt
     hat.</p>
</foliation>
<foliation>
  <p>Folio numbers were added in brown ink by Árni Magnússon ca.
     1720-1730 in the upper right corner of all recto-pages.</p>
</foliation>
```

Where a manuscript contains traces of more than one foliation, each should be recorded as a distinct <foliation> element and optionally given a distinct value for its *xml:id* attribute. The <locus> element discussed in *11.3.5. References to Locations within a Manuscript* can then indicate which foliation scheme is being cited by means of its *scheme* attribute, which points to this identifier:

```
<foliation  xml:id="original">
  <p>Original foliation in red roman numerals in the
     middle of the outer margin of each recto</p>
</foliation>
<foliation  xml:id="modern">
  <p>Foliated in pencil in the top right corner of each
     recto page.</p>
</foliation>
<!-- ... -->
<locus  scheme="#modern">ff 1-20</locus>
```

#### 11.7.1.5 Condition

The <condition> element is used to summarize the overall physical state of a manuscript, in particular where such information is not recorded elsewhere in the description. It should not, however, be used to describe changes or repairs to a manuscript, as these are more appropriately described as a part of its custodial history (see *11.9.1.2. Availability and Custodial History*). It should be supplied within the <supportDesc> element, if it discusses the condition of the physical support of the manuscript; within the <bindingDesc> or <binding> elements (*11.7.3.1. Binding Descriptions*) if it discusses only the condition of the binding or bindings concerned; or within the <sealDesc> element if it discusses the condition of any seal attached to the manuscript.

```
<supportDesc>
  <condition>
    <p>The manuscript shows signs of damage from water and mould on its
        outermost leaves.</p>
  </condition>
</supportDesc>
```

```
<condition>
  <p>Despite tears on many of the leaves the codex is reasonably well
     preserved. The top and the bottom of f. 1 is damaged, and only a thin slip is
     left of the original second leaf (now foliated as 1bis). The lower margin of f.
     92 has been cut away. There is a lacuna of one leaf between ff. 193 and 194. The
     manuscript ends defectively (there are approximately six leaves
     missing).</p>
</condition>
```

#### 11.7.1.6 Layout Description

The second part of the <objectDesc> element is the <layoutDesc> element, which is used to describe and document the *mise-en-page* of the manuscript, that is the way in which text and illumination are arranged on the page, specifying for example the number of written, ruled, or pricked lines and columns per page, size of margins, distinct textual streams such as glosses, commentaries, etc. This may be given as a simple series of paragraphs. Alternatively, one or more different layouts may be identified within a single manuscript, each described by its own <layout> element.

`<layoutDesc>` (layout description) collects the set of layout descriptions applicable to a manuscript or other object.

`<layout>` (layout) describes how text is laid out on the page or surface of the object, including information about any ruling, pricking, or other evidence of page-preparation techniques.

Where the <layout> element is used, the layout will often be sufficiently regular for the attributes on this element to convey all that is necessary; more usually however a more detailed treatment will be required. The attributes are provided as a convenient shorthand for commonly occurring cases, and should not be used except where the layout is regular. The value `NA` (not-applicable) should be used for cases where the layout is either very irregular, or where it cannot be characterized simply in terms of lines and columns, for example, where blocks of commentary and text are arranged in a regular but complex pattern on each page

The following examples indicate the range of possibilities:

```
<layout ruledLines="25 32">
  <p>Most pages have between 25 and 32 long lines ruled in lead.</p>
</layout>
<layout columns="1" writtenLines="24">
  <p>Written in one column throughout; 24 lines per page.</p>
</layout>
<layout>
  <p>Written in 3 columns, with 8 lines of text and interlinear glosses in
     the centre, and up to 26 lines of gloss in the outer two columns. Double
     vertical bounding lines ruled in hard point on hair side. Text lines ruled
     faintly in lead. Remains of prickings in upper, lower, and outer (for 8 lines of
     text only) margins.</p>
</layout>
```

Where multiple <layout> elements are supplied, the scope for each specification can be indicated by means of <locus> elements within the content of the element, as in the following example:

```
<layoutDesc>
  <layout ruledLines="25 32">
    <p>On <locus from="1r" to="202v">fols 1r-200v</locus> and <locus from="210r" to="212v">fols
210r-212v</locus> there are between 25 and 32 ruled lines.</p>
  </layout>
  <layout ruledLines="34 50">
    <p>On <locus from="203r" to="209v">fols 203r-209v</locus> there are between 34
        and 50 ruled lines.</p>
  </layout>
</layoutDesc>
```

### 11.7.2 Writing, Decoration, and Other Notations

The second group of elements within a structured physical description concerns aspects of the writing, illumination, or other notation (notably, music) found in a manuscript, including additions made in later hands—the 'text', as it were, as opposed to the carrier.

`<handDesc>` (description of hands) contains a description of all the different hands used in a manuscript or other object.

<handNote> (note on hand) describes a particular style or hand distinguished within a manuscript.

<scriptDesc> (script description) contains a description of the scripts used in a manuscript or other object.

<scriptNote> describes a particular script distinguished within the description of a manuscript or similar resource.

<typeDesc> (typeface description) contains a description of the typefaces or other aspects of the printing of an incunable or other printed source.

<typeNote> (typographic note) describes a particular font or other significant typographic feature distinguished within the description of a printed resource.

<decoDesc> (decoration description) contains a description of the decoration of a manuscript or other object, either as in paragraphs, or as one or more <decoNote> elements.

<decoNote> (note on decoration) contains a note describing either a decorative component of a manuscript or other object, or a fairly homogenous class of such components.

<musicNotation> (music notation) contains description of type of musical notation.

<additions> (additions) contains a description of any significant additions found within a manuscript or other object, such as marginalia or other annotations.

#### 11.7.2.1 Writing

The <handDesc> element can contain a short description of the general characteristics of the writing observed in a manuscript, as in the following example:

```
<handDesc>
  <p>Written in a <term>late Caroline minuscule</term>; versals in a form of
  <term>rustic capitals</term>; although the marginal and interlinear gloss is
     written in varying shades of ink that are not those of the main text, text and
     gloss appear to have been copied during approximately the same time span.</p>
</handDesc>
```

Note the use of the <term> element to mark specific technical terms within the context of the <handDesc> element.

Where several distinct hands have been identified, this fact can be registered by using the *hands* attribute, as in the following example:

```
<handDesc  hands="2">
  <p>The manuscript is written in two contemporary hands, otherwise unknown, but
     clearly those of practised scribes. Hand I writes ff. 1r-22v and hand II ff. 23
     and 24. Some scholars, notably Verner Dahlerup and Hreinn Benediktsson, have
     argued for a third hand on f. 24, but the evidence for this is
     insubstantial.</p>
</handDesc>
```

Alternatively, or in addition, where more specific information about one or more of the hands identified is to be recorded, the <handNote> element should be used, as in the following example:

```
<handDesc  hands="3">
  <handNote  xml:id="Eirsp-1"  scope="minor">
    <p>The first part of the manuscript, <locus  from="1v"  to="72v:4">fols
           1v-72v:4</locus>, is written in a practised Icelandic Gothic bookhand. This
hand
        is not found elsewhere.</p>
  </handNote>
  <handNote  xml:id="Eirsp-2"  scope="major">
    <p>The second part of the manuscript, <locus  from="72v:4"  to="194v">fols
           72v:4-194</locus>, is written in a hand contemporary with the first; it can
```

```
also
       be found in a fragment of <title>Knýtlinga saga</title>, <ref>AM 20b II
          fol.</ref>.</p>
  </handNote>
  <handNote  xml:id="Eirsp-3"  scope="minor">
    <p>The third hand has written the majority of the chapter headings. This hand
       has been identified as the one also found in <ref>AM 221
          fol.</ref>.</p>
  </handNote>
</handDesc>
```

Note here the use of the <locus> element, discussed in section *11.3.5. References to Locations within a Manuscript*, to specify exactly which parts of a manuscript are written by a given hand.

When a full or partial transcription of a manuscript is available in addition to the manuscript description, the <handShift> element described in *12.3.2.1. Document Hands* can be used to link the relevant parts of the transcription to the appropriate <handNote> element in the description: for example, at the point in the transcript where the second hand listed above starts (i.e. at folio 72v:4), we might insert <handShift new="#Eirsp-2"/>.

Additions, notes, drawings etc. (e.g. <add>, <note> and <figure>) made by other hands in the text, can be linked to the corresponding <handNote> element using the *hand* attribute.

The elements <typeDesc>, and <typeNote> are used to provide information about the printing of a source, in exactly the same way as the <handDesc> or <handNote> elements provide information about its writing. They are provided for the convenience of those using this module to provide information about early printed sources and incunables. The <typeDesc> element can simply provide a summary description:

```
<typeDesc>
  <p>Uses a mixture of Roman and Black Letter types.</p>
</typeDesc>
```

Where detailed information about individual typefaces is to be recorded, this may be done using the <typeNote> element:

```
<typeDesc>
  <summary>Uses a mixture of Roman and Black Letter types.</summary>
  <typeNote>Antiqua typeface, showing influence of Jenson's Venetian
     fonts.</typeNote>
  <typeNote>The black letter face is a variant of Schwabacher.</typeNote>
</typeDesc>
```

Where information is required about both typography and written script, for example where a printed book contains extensive handwritten annotation, both <handDesc> and <typeDesc> elements should be supplied. Similarly, in the following example, the source text is a typescript with extensive handwritten annotation:

```
<typeDesc>
  <typeNote  xml:id="TSET">Authorial typescript, probably produced on Eliot's own
     Remington. </typeNote>
</typeDesc>
<handDesc>
  <handNote  xml:id="EP"  medium="red-ink">Ezra Pound's annotations.</handNote>
  <handNote  xml:id="TSE"  medium="black-ink">Commentary in Eliot's hand.</handNote>
</handDesc>
```

The elements <scriptNote> and <scriptDesc> may be used in exactly the same way to document a script used in this and other manuscripts, for example to record that this script was used mainly for the production of books or

for charters; or that it is characteristic of some geographical area or scriptorium or date. Such information as the letter forms characteristic of this script may also be recorded. By contrast, the <handNote> element would be used to document the way that a particular scribe uses a script, for example with long or short descenders, or using a pen which is cut in a different way, or an ink of a given colour, and so forth.

As with <typeNote>, the <scriptNote> element can be used in combination with <handNote>.

#### 11.7.2.2 Decoration

It can be difficult to draw a clear distinction between aspects of a manuscript which are purely physical and those which form part of its intellectual content. This is particularly true of illuminations and other forms of decoration in a manuscript. We propose the following elements for the purpose of delimiting discussion of these aspects within a manuscript description, and for convenience locate them all within the physical description, despite the fact that the illustrative features of a manuscript will in many cases also be seen as constituting part of its intellectual content.

The <decoDesc> element may contain simply one or more paragraphs summarizing the overall nature of the decorative features of the manuscript, as in the following example:

```
<decoDesc>
  <p>The decoration comprises two full page miniatures, perhaps added by the
     original owner, or slightly later; the original major decoration consists of
     twenty-three large miniatures, illustrating the divisions of the Passion
     narrative and the start of the major texts, and the major divisions of the
     Hours; seventeen smaller miniatures, illustrating the suffrages to saints; and
     seven historiated initials, illustrating the pericopes and major prayers.</p>
</decoDesc>
```

Alternatively, it may contain a series of more specific typed <decoNote> elements, each summarizing a particular aspect or individual instance of the decoration present, for example the use of miniatures, initials (historiated or otherwise), borders, diagrams, etc., as in the following example:

```
<decoDesc>
  <decoNote type="miniature">
    <p>One full-page miniature, facing the beginning of the first Penitential
        Psalm.</p>
  </decoNote>
  <decoNote type="initial">
    <p>One seven-line historiated initial, commencing the first Penitential
        Psalm.</p>
  </decoNote>
  <decoNote type="initial">
    <p>Six four-line decorated initials, commencing the second through the seventh
        Penitential Psalm.</p>
  </decoNote>
  <decoNote type="initial">
    <p>Some three hundred two-line versal initials with pen-flourishes, commencing
        the psalm verses.</p>
  </decoNote>
  <decoNote type="border">
    <p>Four-sided border decoration surrounding the miniatures and three-sided
        border decoration accompanying the historiated and decorated initials.</p>
  </decoNote>
</decoDesc>
```

Where more exact indexing of the decorative content of a manuscript is required, the standard TEI elements <term> or <index> may be used within the prose description to supply or delimit appropriate iconographic terms, as in the following example:

```
<decoDesc>
  <decoNote  type="miniatures">
    <p>Fourteen large miniatures with arched tops, above five lines of text: <list>
       <item>
         <locus>fol. 14r</locus>Pericopes. <term>St. John writing on Patmos</term>,
               with the Eagle holding his ink-pot and pen-case; some flaking of
pigment,
               especially in the sky</item>
       <item>
         <locus>fol. 26r</locus>Hours of the Virgin, Matins.
       <term>Annunciation</term>; Gabriel and the Dove to the right</item>
       <item>
         <locus>fol. 60r</locus>Prime. <term>Nativity</term>; the <term>Virgin and
                  Joseph adoring the Child</term>
       </item>
       <item>
         <locus>fol. 66r</locus>Terce. <term>Annunciation to the Shepherds</term>,
               one with <term>bagpipes</term>
       </item>
<!-- ... -->
     </list>
    </p>
  </decoNote>
</decoDesc>
```

#### 11.7.2.3 Musical Notation

Where a manuscript contains music, the <musicNotation> element may be used to describe the form of notation employed, as in the following examples:

```
<musicNotation>
  <p>Square notation on 4-line red staves.</p>
</musicNotation>
```

```
<musicNotation>
  <p>Neumes in campo aperto of the St. Gall type.</p>
</musicNotation>
```

If a manuscript employs more than one notation, they must both be described within the same <musicNotation> element, for example as different list items.

#### 11.7.2.4 Additions and Marginalia

The <additions> element can be used to list or describe any additions to the manuscript, such as marginalia, scribblings, doodles, etc., which are considered to be of interest or importance. Such topics may also be discussed or referenced elsewhere in a description, for example in the <history> element, in cases where the marginalia provide evidence of ownership. Note that this element may not be repeated within a single manuscript description. If several different kinds of additional matter are discussed, the content may be structured as a labelled list or a series of paragraphs. Some examples follow:

```
<additions>
  <p>Doodles on most leaves, possibly by children, and often quite
     amusing.</p>
</additions>
```

```
<additions>
  <p  xml:lang="fr">Quelques annotations marginales des XVIe et XVIIe s.</p>
</additions>
```

```
<additions>
  <p>The text of this manuscript is not interpolated with sentences from Royal
     decrees promulgated in 1294, 1305 and 1314. In the margins, however, another
     somewhat later scribe has added the relevant paragraphs of these decrees, see
     pp. 8, 24, 44, 47 etc.</p>
  <p>As a humorous gesture the scribe in one opening of the manuscript, pp. 36 and
     37, has prolonged the lower stems of one letter f and five letters þ and has
     them drizzle down the margin.</p>
</additions>
```

```
<additions>
  <p>Spaces for initials and chapter headings were left by the scribe but not
     filled in. A later, probably fifteenth-century, hand has added initials and
     chapter headings in greenish-coloured ink on fols <locus>8r</locus>,
  <locus>8v</locus>, <locus>9r</locus>, <locus>10r</locus> and <locus>11r</locus>.
     Although a few of these chapter headings are now rather difficult to read, most
     can be made out, e.g. fol. <locus>8rb</locus>
    <quote  xml:lang="is">floti ast<ex>ri</ex>d<ex>ar</ex>
    </quote>; fol.
  <locus>9rb</locus>
    <quote  xml:lang="is">v<ex>m</ex> olaf conung</quote>, and fol.
  <locus>10ra</locus>
    <quote  xml:lang="is">Gipti<ex>n</ex>g ol<ex>a</ex>fs
        k<ex>onun</ex>gs</quote>.</p>
  <p>The manuscript contains the following marginalia: <list>
      <item>Fol. <locus>4v</locus>, left margin: <quote  xml:lang="is">hialmadr
        <ex>ok</ex>
          <lb/>brynjadr</quote>, in a fifteenth-century hand, imitating an addition made
            to the text by the scribe at this point.</item>
      <item>Fol. <locus>5r</locus>, lower margin: <quote  xml:lang="is">þ<ex>e</ex>tta
                þiki m<ex>er</ex> v<ex>er</ex>a gott blek en<ex>n</ex>da kan<ex>n</ex>
ek icki
                betr sia</quote>, in a fifteenth-century hand, probably the same as that
on the
            previous page.</item>
      <item>Fol. <locus>9v</locus>, bottom margin: <quote  xml:lang="is">þessa bok
                uilda eg <sic>gæt</sic> lært med
        <lb/>an Gud gefe myer Gott ad
        <lb/>læra</quote>; seventeenth-century hand.</item>
    </list>
  </p>
  <p>There are in addition a number of illegible scribbles in a later hand (or
     hands) on fols <locus>2r</locus>, <locus>3r</locus>, <locus>5v</locus> and
  <locus>19r</locus>.</p>
</additions>
```

### 11.7.3 Bindings, Seals, and Additional Material

The third major component of the physical description relates to supporting but distinct physical components, such as bindings, seals and accompanying material. These may be described using the following specialist elements:

`<bindingDesc>` (binding description) describes the present and former bindings of a manuscript or other object, either as a series of paragraphs or as a series of distinct <binding> elements, one for each binding of the manuscript.

`<binding>` (binding) contains a description of one binding, i.e. type of covering, boards, etc. applied to a manuscript or other object.

`<sealDesc>` (seal description) describes the seals or similar items related to the object described, either as a series of paragraphs or as a series of <seal> elements.

`<seal>` (seal) contains a description of one seal or similar applied to the object described.

`<accMat>` (accompanying material) contains details of any significant additional material which may be closely associated with the manuscript or object being described, such as non-contemporaneous documents or fragments bound in with it at some earlier historical period.

#### 11.7.3.1 Binding Descriptions

The <bindingDesc> element contains a description of the state of the present and former bindings of a manuscript, including information about its material, any distinctive marks, and provenance information. This may be given as a series of paragraphs if only one binding is being described, or as a series of distinct <binding> elements, each describing a distinct binding where these are separately described. For example:

```
<bindingDesc>
  <p>Sewing not visible; tightly rebound over 19th-century pasteboards, reusing
     panels of 16th-century brown leather with gilt tooling à la fanfare, Paris c.
     1580-90, the centre of each cover inlaid with a 17th-century oval medallion of
     red morocco tooled in gilt (perhaps replacing the identifying mark of a previous
     owner); the spine similarly tooled, without raised bands or title-piece;
     coloured endbands; the edges of the leaves and boards gilt. Boxed.</p>
</bindingDesc>
```

Within a binding description, the elements <decoNote> and <condition> are available, as alternatives to <p>, for paragraphs dealing exclusively with information about decorative features of a binding, or about its condition, respectively.

```
<binding>
  <p>Bound, s. XVIII (?), in <material>diced russia leather</material>
     retaining most of the original 15th century metal ornaments (but with some
     replacements) as well as the heavy wooden boards.</p>
  <decoNote>
    <p>On each cover: alternating circular stamps of the Holy Monogram, a
        sunburst, and a flower.</p>
  </decoNote>
  <decoNote>
    <p>On the cornerpieces, one of which is missing, a rectangular stamp
        of the Agnus Dei.</p>
  </decoNote>
  <condition>Front and back leather inlaid panels very badly worn.</condition>
  <p>Rebacked during the 19th century.</p>
</binding>
```

As noted above, (*11.7.1.5. Condition*) the element <condition> may also be used as an alternative to <p> for paragraphs concerned exclusively with the condition of a binding, where this has not been supplied as part of the physical description.

#### 11.7.3.2 Seals

The <sealDesc> element supplies information about the seal(s) attached to documents to guarantee their integrity, or to show authentication of the issuer or consent of the participants. It may contain one or more paragraphs summarizing the overall nature of the seals, or may contain one or more <seal> elements.

```
<sealDesc>
  <seal  n="1"  type="pendant"
    subtype="cauda_duplex">
    <p>Round seal of <name>Anders Olufsen</name> in black wax: <bibl>
       <ref>DAS
              930</ref>
     </bibl>. Parchment tag, on which is written: <quote>pertinere nos
           predictorum placiti nostri iusticarii precessorum dif</quote>.</p>
  </seal>
  <seal  n="2"  type="pendant"
    subtype="cauda_duplex">
    <p>The seal of <name>Jens Olufsen</name> in black wax: <bibl>
       <ref>DAS
              1061</ref>
     </bibl>. Legend: <quote>S IOHANNES OLAVI</quote>. Parchment tag on
        which is written: <quote>Woldorp Iohanne G</quote>.</p>
  </seal>
</sealDesc>
```

#### 11.7.3.3 Accompanying Material

The circumstance may arise where material not originally part of a manuscript is bound into or otherwise kept with a manuscript. In some cases this material would best be treated in a separate <msPart> element (see *11.10. Manuscript Parts* below). There are, however, cases where the additional matter is not self-evidently a distinct manuscript: it might, for example, be a set of notes by a later scholar, or a file of correspondence relating to the manuscript. The <accMat> element is provided as a holder for this kind of information.

`<accMat>` (accompanying material) contains details of any significant additional material which may be closely associated with the manuscript or object being described, such as non-contemporaneous documents or fragments bound in with it at some earlier historical period.

Here is an example of the use of this element, describing a note by the Icelandic manuscript collector Árni Magnússon which has been bound with the manuscript:

```
<accMat>
  <p>A slip in Árni Magnússon's hand has been stuck to the pastedown on the inside
     front cover; the text reads: <quote  xml:lang="is">Þidreks Søgu þessa hefi eg
        feiged af Sekreterer Wielandt Anno 1715 i Kaupmanna høfn. Hun er, sem eg sie,
        Copia af Austfirda bókinni (Eidagás) en<ex>n</ex> ecki progenies Brædratungu
        bokarinnar. Og er þar fyrer eigi i allan<ex>n</ex> máta samhlioda
        þ<ex>eir</ex>re er Sr Jon Erlendz son hefer ritad fyrer Mag. Bryniolf. Þesse
        Þidreks Saga mun vera komin fra Sr Vigfuse á Helgafelle.</quote>
  </p>
</accMat>
```

## 11.8 History

The following elements are used to record information about the history of a manuscript:

`<history>` (history) groups elements describing the full history of a manuscript, manuscript part, or other object.

`<origin>` (origin) contains any descriptive or other information concerning the origin of a manuscript, manuscript part, or other object.

`<provenance>` (provenance) contains any descriptive or other information concerning a single identifiable episode during the history of a manuscript, manuscript part, or other object after its creation but before its acquisition.

`<acquisition>` (acquisition) contains any descriptive or other information concerning the process by which a manuscript or manuscript part or other object entered the holding institution.

The three components of the <history> element all have the same substructure, consisting of one or more paragraphs marked as <p> elements. Each of these three elements is also a member of the `att.datable` attribute class, itself a member of the `att.datable.w3c` class, and thus also carries the following optional attributes:

`att.datable.w3c` provides attributes for normalization of elements that contain datable events conforming to the W3C XML Schema Part 2: Datatypes Second Edition.

- *@notBefore* specifies the earliest possible date for the event in standard form, e.g. yyyy-mm-dd.
- *@notAfter* specifies the latest possible date for the event in standard form, e.g. yyyy-mm-dd.

Information about the origins of the manuscript, its place and date of writing, should be given as one or more paragraphs contained by a single <origin> element; following this, any available information on distinct stages in the history of the manuscript before its acquisition by its current holding institution should be included as paragraphs within one or more <provenance> elements. Finally, any information specific to the means by which the manuscript was acquired by its present owners should be given as paragraphs within the <acquisition> element.

Here is a fairly simple example of the use of this element:

```
<history>
  <origin>
    <p>Written in <origPlace>Durham</origPlace> during <origDate  notBefore="1125"
       notAfter="1175">the mid-twelfth
           century</origDate>.</p>
  </origin>
  <provenance>
    <p>Recorded in two medieval catalogues of the books belonging to
    <name  type="org">Durham Priory</name>, made in <date>1391</date> and
    <date>1405</date>.</p>
  </provenance>
  <provenance>
    <p>Given to <name  type="person">W. Olleyf</name> by <name  type="person">William
Ebchester, Prior (1446-56)</name> and later belonged to
    <name  type="person">Henry Dalton</name>, Prior of Holy Island (<name  type="place">Lindisfarne
according to inscriptions on ff. 4v and 5.</p>
  </provenance>
  <acquisition>
    <p>Presented to <name  type="org">Trinity College</name> in
    <date>1738</date> by <name  type="person">Thomas Gale</name> and his son
<name  type="person">Roger</name>.</p>
  </acquisition>
</history>
```

Here is a fuller example, demonstrating the use of multiple <provenance> elements where distinct periods of ownership for the manuscript have been identified:

```
<history>
  <origin  notBefore="1225"  notAfter="1275"> Written in Spain or Portugal in the
     middle of the 13th century (the date 1042, given in a marginal note on f. 97v,
     cannot be correct.)</origin>
  <provenance>The Spanish scholar <name  type="person">Benito Arias Montano</name>
     (1527-1598) has written his name on f. 97r, and may be presumed to have owned
     the manuscript. </provenance>
  <provenance>It came somehow into the possession of <foreign  xml:lang="da">etatsråd</foreign>
    <name  type="person">Holger Parsberg</name> (1636-1692), who has written his name
     twice, once on the front pastedown and once on f. 1r, the former dated
  <date>1680</date> and the latter <date>1682</date>.</provenance>
  <provenance>Following Parsberg's death the manuscript was bought by
```

```
  <foreign>etatsråd</foreign>
    <name  type="person">Jens Rosenkrantz</name> (1640-1695) when Parsberg's library
     was auctioned off (23 October 1693).</provenance>
  <acquisition  notBefore="1696"
    notAfter="1697">The manuscript was acquired by
     Árni Magnússon from the estate of Jens Rosenkrantz, presumably at auction (the
     auction lot number 468 is written in red chalk on the flyleaf), either in 1696
     or 97.</acquisition>
</history>
```

## 11.9 Additional Information

Three categories of additional information are provided for by the scheme described here, grouped together within the <additional> element described in this section.

`<adminInfo>` (administrative information) contains information about the present custody and availability of the manuscript or other object, and also about the record description itself.

`<surrogates>` (surrogates) contains information about any representations of the manuscript or other object being described which may exist in the holding institution or elsewhere.

`<listBibl>` (citation list) contains a list of bibliographic citations of any kind.

None of these specialized constituent elements of <additional> is required. If any is supplied, it may appear once only; furthermore, the order in which elements are supplied should be as specified above. Alternatively, additional information may be provided as either one or more paragraphs or anonymous blocks, marked up as a series of <p> or <ab> elements.

### 11.9.1 Administrative Information

The <adminInfo> element is used to hold information relating to the curation and management of a manuscript. This may be supplied as a note using the global <note> element. Alternatively, different aspects of this information may be presented grouped within one of the following specialized elements:

`<recordHist>` (recorded history) provides information about the source and revision status of the parent manuscript or object description itself.

`<availability>` (availability) supplies information about the availability of a text, for example any restrictions on its use or distribution, its copyright status, any licence applying to it, etc.

`<custodialHist>` (custodial history) contains a description of a manuscript or other object's custodial history, either as running prose or as a series of dated custodial events.

#### 11.9.1.1 Record History

The <recordHist> element may contain simply a series of paragraphs. Alternatively it may contain a <source> element, followed by an optional series of <change> elements.

`<source>` (source) describes the original source for the information contained with a manuscript or object description.

`<change>` (change) documents a change or set of changes made during the production of a source document, or during the revision of an electronic file.

The <source> element is used to document the primary source of information for the record containing it, in a similar way to the standard TEI <sourceDesc> element within a TEI Header. If the record is a new one, made without reference to anything other than the manuscript itself, then it may simply contain a <p> element, as in the following example:

```
<source>
  <p>Directly catalogued from the original manuscript.</p>
</source>
```

Frequently, however, the record will be derived from some previously existing description, which may be specified using the standard TEI <bibl> element, as in the following example:

```
<recordHist>
  <source>
    <p>Information transcribed from <bibl>
       <title>The index of Middle English
              verse</title>
       <biblScope>123</biblScope>
     </bibl>.</p>
  </source>
</recordHist>
```

If, as is likely, a full bibliographic description of the source from which cataloguing information was taken is included within the <listBibl> element contained by the current <additional> element, or elsewhere in the current document, then it need not be repeated here. Instead, it should be referenced using the standard TEI <ref> element, as in the following example:

```
<additional>
  <adminInfo>
    <recordHist>
      <source>
        <p>Information transcribed from <bibl>
           <ref  target="#IMEV">IMEV</ref>
                 123</bibl>.</p>
      </source>
    </recordHist>
  </adminInfo>
  <listBibl>
    <bibl  xml:id="IMEV">
      <author>Carleton Brown</author> and <author>Rossell Hope Robbins</author>
      <title  level="m">The index of Middle English verse</title>
      <pubPlace>New York</pubPlace>
      <date>1943</date>
    </bibl>
<!-- other bibliographic records relating to this manuscript here -->
  </listBibl>
</additional>
```

The <change> element may also appear within the <revisionDesc> element of the standard TEI header; its use here is intended to signal the similarity of function between the two container elements. Where the TEI header should be used to document the revision history of the whole electronic file to which it is prefixed, the <recordHist> element may be used to document changes at a lower level, relating to the individual description, as in the following example:

```
<change  when="2005-03-10">On 10 March 2005 <name>MJD</name> added provenance
information</change>
```

#### 11.9.1.2 Availability and Custodial History

The <availability> element is another element also available in the TEI header, which should be used here to supply any information concerning access to the current manuscript, such as its physical location (where this is not implicit in its identifier), any restrictions on access, information about copyright, etc.

```
<availability>
  <p>Viewed by appointment only, to be arranged with curator.</p>
```

```
</availability>
<availability>
  <p>In conservation, Jan. - Mar., 2002. On loan to the Bayerische
     Staatsbibliothek, April - July, 2002.</p>
</availability>
<availability>
  <p>The manuscript is in poor condition, due to many of the leaves being brittle
     and fragile and the poor quality of a number of earlier repairs; it should
     therefore not be used or lent out until it has been conserved.</p>
</availability>
```

The <custodialHist> record is used to describe the custodial history of a manuscript, recording any significant events noted during the period that it has been located within its holding institution. It may contain either a series of <p> elements, or a series of <custEvent> elements, each describing a distinct incident or event, further specified by a *type* attribute, and carrying dating information by virtue of its membership in the `att.datable` class, as noted above.

`<custEvent>` (custodial event) describes a single event during the custodial history of a manuscript or other object.

Here is an example of the use of this element:

```
<custodialHist>
  <custEvent  type="conservation"
    notBefore="1961-03-01"  notAfter="1963-02-28">
    <p>Conserved between March 1961 and February 1963 at Birgitte Dalls
        Konserveringsværksted.</p>
  </custEvent>
  <custEvent  type="photography"
    notBefore="1988-05-01"  notAfter="1988-05-30">
    <p>Photographed in May 1988 by AMI/FA.</p>
  </custEvent>
  <custEvent  type="transfer"
    notBefore="1989-11-13"  notAfter="1989-11-13">
    <p>Dispatched to Iceland 13 November 1989.</p>
  </custEvent>
</custodialHist>
```

## 11.9.2 Surrogates

The <surrogates> element is used to provide information about representations such as photographs or other representations of the manuscript which may exist within the holding institution or elsewhere.

`<surrogates>` (surrogates) contains information about any representations of the manuscript or other object being described which may exist in the holding institution or elsewhere.

The <surrogates> element should not be used to repeat information about representations of the manuscript available within published works; this should normally be documented within the <listBibl> element within the <additional> element. However, it is often also convenient to record information such as negative numbers or digital identifiers for unpublished collections of manuscript images maintained within the holding institution, as well as to provide more detailed descriptive information about the surrogate itself. Such information may be provided as prose paragraphs, within which identifying information about particular surrogates may be presented using the standard TEI <bibl> element, as in the following example:

```
<surrogates>
  <bibl>
    <title  type="gmd">microfilm (master)</title>
```

```
    <idno>G.neg. 160</idno>
     n.d.</bibl>
  <bibl>
    <title  type="gmd">microfilm (archive)</title>
    <idno>G.pos. 186</idno>
     n.d.</bibl>
  <bibl>
    <title  type="gmd">b/w prints</title>
    <idno>AM 795 4to</idno>
    <date  when="1999-01-27">27 January 1999</date>
    <note>copy of G.pos.
        186</note>
  </bibl>
  <bibl>
    <title  type="gmd">b/w prints</title>
    <idno>reg.nr. 75</idno>
    <date  when="1999-01-25">25 January 1999</date>
    <note>photographs of the spine, outside covers, stitching etc.</note>
  </bibl>
</surrogates>
```

Note the use of the specialized form of title (*general material designation*) to specify the kind of surrogate being documented.

At a later revision, the content of the <surrogates> element is likely to be expanded to include elements more specifically intended to provide detailed information such as technical details of the process by which a digital or photographic image was made. For information about the inclusion of digital facsimile images within a TEI document, refer also to *12.1. Digital Facsimiles*.

## 11.10 Manuscript Parts

The <msPart> element may be used in cases where manuscripts or parts of manuscripts that were originally physically separate have been bound together and/or share the same call number.

`<msPart>` (manuscript part) contains information about an originally distinct manuscript or part of a manuscript, which is now part of a composite manuscript.

Since each component of such a composite manuscript will in all likelihood have its own content, physical description, history, and so on, the structure of <msPart> is in the main identical to that of <msDesc>, allowing one to retain the top level of identity (<msIdentifier>), but to branch out thereafter into as many parts, or even subparts, as necessary.

```
<msDesc  xml:id="KBR_ms_10066-77"
  xml:lang="en"  type="composite">
  <msIdentifier>
    <settlement  key="tgn_7007868">Brussels</settlement>
    <repository>Koninklijke Bibliotheek van België / Bibliothèque royale de
        Belgique</repository>
    <idno>ms. 10066-77</idno>
  </msIdentifier>
  <msContents>
    <summary  xml:lang="la">Miscellany of various texts; Prudentius,
        Psychomachia; Physiologus de natura animantium</summary>
    <textLang  mainLang="la">Latin</textLang>
  </msContents>
  <physDesc>
    <objectDesc  form="composite_manuscript"/>
  </physDesc>
```

```
  <msPart>
    <msIdentifier>
      <idno>ms. 10066-77 ff. 140r-156v</idno>
    </msIdentifier>
    <msContents>
      <summary xml:lang="la">Physiologus</summary>
      <textLang mainLang="la">Latin</textLang>
    </msContents>
  </msPart>
  <msPart>
    <msIdentifier>
      <idno>ms. 10066-77 ff. 112r-139r</idno>
    </msIdentifier>
    <msContents>
      <summary xml:lang="la">Prudentius, Psychomachia</summary>
      <textLang mainLang="la">Latin</textLang>
    </msContents>
  </msPart>
</msDesc>
```

## 11.11 Manuscript Fragments

The <msFrag> element may be used inside <msDesc> when encoding one or more fragments of a scattered or fragmented manuscript. The fragment(s) described in a single <msDesc> element may be held either at several institutions or at a single institution, so different call numbers may be attached to the fragments. Inside the <msFrag> element, information about the single fragment or each dispersed part is provided: e.g. the current shelfmark or call number, the labels of the range of folios concerned if the fragment currently forms part of a larger manuscript, dimensions, extent, title, author, annotations, illuminations and so on.

`<msFrag>` (manuscript fragment) contains information about a fragment described in relation to a prior context, typically as a description of a virtual reconstruction of a manuscript or other object whose fragments were catalogued separately.

One well-known example of fragmentation is the Old Church Slavonic manuscript known as *Codex Suprasliensis*, substantial parts of which are to be found in three separate repositories, in Ljubljana, Warsaw, and St. Petersburg. This manuscript should be represented using one single <msDesc> element in which <msName> is used to identify the fragmented manuscript, along with three distinct <msFrag> elements, each of which contains the current identifier of one of the fragments, a physical description, and other related information:

```
<msDesc>
  <msIdentifier>
    <msName xml:lang="la">Codex Suprasliensis</msName>
  </msIdentifier>
  <msFrag>
    <msIdentifier>
      <settlement>Ljubljana</settlement>
      <repository>Narodna in univerzitetna knjiznica</repository>
      <idno>MS Kopitar 2</idno>
    </msIdentifier>
    <msContents>
      <summary>Contains ff. 10 to 42 only</summary>
    </msContents>
  </msFrag>
  <msFrag>
    <msIdentifier>
      <settlement>Warszawa</settlement>
      <repository>Biblioteka Narodowa</repository>
```

```
      <idno>BO 3.201</idno>
    </msIdentifier>
  </msFrag>
  <msFrag>
    <msIdentifier>
      <settlement>Sankt-Peterburg</settlement>
      <repository>Rossiiskaia natsional'naia biblioteka</repository>
      <idno>Q.p.I.72</idno>
    </msIdentifier>
  </msFrag>
</msDesc>
```

## 11.12 Module for Manuscript Description

The module described in this chapter makes available the following components:

**Module msdescription: Manuscript Description**

- *Elements defined*: accMat acquisition additional additions adminInfo altIdentifier binding bindingDesc catchwords collation collection colophon condition custEvent custodialHist decoDesc decoNote depth dim dimensions explicit filiation finalRubric foliation handDesc height heraldry history incipit institution layout layoutDesc locus locusGrp material msContents msDesc msFrag msIdentifier msItem msItemStruct msName msPart musicNotation objectDesc objectType origDate origPlace origin physDesc provenance recordHist repository rubric scriptDesc seal sealDesc secFol signatures source stamp summary support supportDesc surrogates typeDesc typeNote watermark width
- *Classes defined*: att.msClass att.msExcerpt model.physDescPart

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 12 Representation of Primary Sources

This chapter describes elements that may be used to represent primary source materials, such as manuscripts, printed books, ephemera, or other textual documents. Some of these specialized elements, particularly at phrase-level, add to the other elements available within <text> to deal with textual phenomena more specific to primary source transcription. Other structural and block-level elements described here can be used to represent primary source materials by prioritizing the encoding of their spatial features over their logical textual structure (that is, the elements described in chapter *4. Default Text Structure*). These elements, <facsimile>, <sourceDoc>, and their children, may be used in parallel and in combination with an encoding of logical text structures with <text>, or as standalone representations. The element <sourceDoc> in particular provides a way of combining facsimile and transcriptions by embedding transcribed text. This approach focuses on physical and textual features that can be primarily described spatially, such as the sequence of pages in a manuscript, or the layout of a printed page. This is not meant to be the only way of transcribing primary sources in TEI, or even a preferred way; which approach is more appropriate will depend on the specific needs of your project.

Although this chapter discusses manuscript materials more frequently than other forms of written text, most of the recommendations presented are equally applicable to facsimiles of a wide variety of media, including printed matter, monumental inscriptions, and art. Each medium has its own vocabulary of agents. In the following examples, terms such as 'scribe', 'author', 'editor', 'annotator' or 'corrector' may be re-interpreted in terms more appropriate to the medium being transcribed. In printed material, for example, the 'compositor' plays a role analogous to the 'scribe', while in an authorial manuscript, the 'author' and the 'scribe' are the same person.

This module may be used in conjunction with other modules. These recommendations are not intended to meet every transcriptional circumstance likely to be faced by any scholar. They are intended rather as a base to enable encoding of the most common phenomena found in the course of scholarly transcription of primary source materials. These guidelines do not address the encoding of physical description of textual witnesses: the materials of the carrier, the medium of the inscribing implement, the organisation of the carrier materials themselves (such as quiring, collation), authorial instructions or scribal markup, etc., except insofar as these are involved in the broader question of manuscript description, as addressed by the `msdescription` module described in chapter *11. Manuscript Description*.

This chapter begins by describing elements for handling digitally-encoded images of primary source materials for the purpose of creating digital facsimiles using the <facsimile> element (*12.1. Digital Facsimiles*).

The next section (*12.2. Combining Transcription with Facsimile*) describes two ways of combining a facsimile images with a transcription; either by referencing a parallel transcription in <text>, or by providing an 'embedded' transcription that prioritizes the encoding of a resource's spatial features via the <sourceDoc> element and a number of transcriptional elements.

Section *12.3. Scope of Transcriptions* documents elements that support scholars in recording information about specific features of the text written on its physical carrier, such as *12.3.1. Altered, Corrected, and Erroneous Texts* and *12.3.2. Hands and Responsibility*

Section *12.4. Aspects of Layout* describes how complex page layouts may be represented.

Section *12.6. Headers, Footers, and Similar Matter* introduces the element <fw> (*forme work*) for encoding material repeated from page to page that falls outside the stream of the text.

Section *12.7. Identifying Changes and Revisions* describes how to document changes made during the production or revision of a primary source.

The chapter concludes with a technical overview of the structure and organization of the module described here. Some elements from other chapters are recontexualized for situations involving the transcription of primary source materials, whether within <text> or <sourceDoc>. Therefore, this overview should be read in conjunction with chapters *3. Elements Available in All TEI Documents* and *5. Characters, Glyphs, and Writing Modes*.

## 12.1 Digital Facsimiles

A common approach in the TEI to representing pre-existing sources involves transcribing or otherwise converting sources into character form before marking them up. However, it is also a common practice to make a different form of 'digital text' that is instead composed of digital images of the original source, typically one per page, or other written surface. We call such a resource a *digital facsimile*. A digital facsimile may, in the simplest case, just consist of a collection of images, with some metadata to identify them and the source materials portrayed. It may sometimes contain a variety of images of the same source pages, perhaps of different resolutions, or of different

kinds. Such a collection may form part of any kind of document, for example a commentary of a codicological or paleographic nature, where there is a need to align explanatory text with image data. It may also be complemented by a transcribed or encoded version of the original source, which may be linked to the page images or 'embedded' as discussed in *12.2. Combining Transcription with Facsimile*. In this section we present elements designed to support these various possibilities and discuss the associated mechanisms provided by these Guidelines.

When this module is included in a schema, the class `att.global` is extended to include two new pointer attributes, *facs* and *change*:

`att.global.facs` provides attributes used to express correspondence between an element and all or part of a facsimile image or surface.
- `@facs` (facsimile) points to one or more images, portions of an image, or surfaces which correspond to the current element.

`att.global.change` provides attributes allowing its member elements to specify one or more states or revision campaigns with which they are associated.
- `@change` points to one or more <change> elements documenting a state or revision campaign to which the element bearing this attribute and its children have been assigned by the encoder.

The *change* attribute is discussed further below in section *12.7. Identifying Changes and Revisions*. The *facs* attribute is used to associate any element in a transcription with an image of the corresponding part of the source, by means of the usual URI pointing mechanism.

In the simple case where a digital text is composed of page images, the *facs* attribute on the <pb> element may be used to associate each image with an appropriate point in the text:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!--...-->
  </teiHeader>
  <text>
    <pb  facs="page1.png"/>
<!-- text contained on page 1 is encoded here -->
    <pb  facs="page2.png"/>
<!-- text contained on page 2 is encoded here -->
  </text>
</TEI>
```

By convention, this encoding indicates that the image indicated by the *facs* attribute represents the whole of the text following the <pb> (page beginning) element, up to the next <pb> element. Any convenient milestone element (see further *3.11.3. Milestone Elements*) could be used in the same way; for example if the images represent individual columns, the <cb> element might be used. Though simple, this method has some drawbacks. It does not scale well to more complex cases where, for example, the images do not correspond exactly with transcribed pages, or where the intention is to align specific marked up elements with detailed images, or parts of images. The management of information about the images may become more difficult if references to them are scattered through many files rather than being concentrated in a single identifiable location. Nevertheless, this solution may be adequate for many straightforward 'digital library' applications.

The recommended approach to encoding facsimiles is instead to use the *facs* attribute in conjunction with the elements <facsimile> or <sourceDoc>, and the elements <surface>, <surfaceGrp>, and <zone>, which are also provided by this module. These elements make it possible to accommodate multiple images of each page, as well as to record the position and relative size of elements identified on any kind of written surface and to link such elements with digital facsimile images of them. Typical applications include the provision of full text search in 'digital facsimile editions', and ways of annotating graphics, for example so as to identify individuals appearing in group portraits and link them to data about the people represented.

The following elements are available to represent components of a digital facsimile:

`<facsimile>` contains a representation of some written source in the form of a set of images rather than as transcribed or encoded text.

`<sourceDoc>` contains a transcription or other representation of a single source document potentially forming part of a *dossier génétique* or collection of sources.

`<surface>` defines a written surface as a two-dimensional coordinate space, optionally grouping one or more graphic representations of that space, zones of interest within that space, and, when using an embedded transcription approach, transcriptions of the writing within them.

`<surfaceGrp>` (surface group) defines any kind of useful grouping of written surfaces, for example the recto and verso of a single leaf, which the encoder wishes to treat as a single unit.

`<zone>` defines any two-dimensional area within a <surface> element.

`@points [att.coordinated]` identifies a two dimensional area by means of a series of pairs of numbers, each of which gives the x,y coordinates of a point on a line enclosing the area.

`<path>` (path) defines any line passing through two or more points within a <surface> element.

`@points` identifies a line within the container or bounding box specified by the parent element by means of a series of two or more pairs of numbers, each of which gives the x,y coordinates of a point on the line.

Either of the <facsimile> or <sourceDoc> elements may be used to represent a digital facsimile. Either may appear within a TEI document along with, or instead of, the <text> element introduced in section *4. Default Text Structure*. The <sourceDoc> element is used when a digital facsimile contains a transcription that prioritizes the encoding of the spatial features and layout of a text document over the text's logical textual structure; the <text> element should be used when it contains a textual transcription focused on its logical features. When the digital facsimile contains only images, however, only <facsimile> elements should be used. In this section, we first discuss the simpler case, returning to the use of the <sourceDoc> element in section *12.2. Combining Transcription with Facsimile* below. When this module is selected therefore, a legal TEI document may thus comprise any of the following:

- a TEI header and a text element
- a TEI header and a facsimile element
- a TEI header and a sourceDoc element
- a TEI header, a facsimile element, and a text element
- a TEI header, one or more sourceDoc or facsimile elements, and a text element

Like the <text> element, a <facsimile> element may also contain an optional <front> or <back> element, used in the same way as described in sections *4.5. Front Matter* and *4.7. Back Matter*.

In the simplest case, a facsimile just contains a series of <graphic> elements, each of which identifies an image file:

```
<facsimile>
  <graphic url="page1.png"/>
  <graphic url="page2.png"/>
  <graphic url="page3.png"/>
  <graphic url="page4.png"/>
</facsimile>
```

If desired, the <binaryObject> element described in *3.10. Graphics and Other Non-textual Components* (or any other element from the `model.graphicLike` class) can be used instead of a <graphic>.

In this simple case, the four page images are understood to represent the complete facsimile, and are to be read in the sequence given. Suppose, however, that the second page of this particular work is available both as an ordinary photograph and as an infra-red image, or in two different resolutions. The <surface> element may be used to group the two image files, since these correspond with the same area of the work:

```
<facsimile>
  <graphic url="page1.png"/>
  <surface>
    <graphic url="page2-highRes.png"/>
    <graphic url="page2-lowRes.png"/>
```

```
  </surface>
  <graphic  url="page3.png"/>
  <graphic  url="page4.png"/>
</facsimile>
```

The <surface> element provides a way of indicating that the two images of page2 represent the same surface within the source material. A *surface* might be one side of a piece of paper or parchment, an opening in a codex treated as a single surface by the writer, a face of a monument, a billboard, a membrane of a scroll, or indeed any two-dimensional surface, of any size.

The <surfaceGrp> element may be used to indicate that two (or more) surfaces are associated in some way, for example because they represent the recto and verso of the same leaf, as in this example:

```
<facsimile>
  <surfaceGrp  n="leaf1">
    <surface>
      <graphic  url="page1.png"/>
    </surface>
    <surface>
      <graphic  url="page2-highRes.png"/>
      <graphic  url="page2-lowRes.png"/>
    </surface>
  </surfaceGrp>
</facsimile>
```

The <surfaceGrp> element may also be useful as a means of identifying other groups of written surfaces, such as adjacent faces of a monument, or gatherings of leaves.

Simply grouping related graphics is not however the main purpose of the <surface> element: rather it is to help identify the location and size of the various two-dimensional spaces constituting the digital facsimile. Note that the actual dimensions of the object represented are not provided by the <surface> element ; rather, the <surface> element defines an abstract coordinate space which may be used to address parts of the image. Four attributes supplied by the `att.coordinated` class are used to define this space.

`att.coordinated` provides attributes that can be used to position their parent element within a two dimensional coordinate system.

- `@ulx` gives the x coordinate value for the upper left corner of a rectangular space.
- `@uly` gives the y coordinate value for the upper left corner of a rectangular space.
- `@lrx` gives the x coordinate value for the lower right corner of a rectangular space.
- `@lry` gives the y coordinate value for the lower right corner of a rectangular space.

By default, the same coordinate space is used for a <surface> and for all of its child elements.[76] It may be most convenient to derive a coordinate space from a digital image of the surface in question such that each pixel in the image corresponds with a whole number of units (typically 1) in the coordinate space. In other cases it may be more convenient to use units such as millimetres. Neither practice implies any specific mapping between the coordinate system used and the actual dimensions of the physical object represented.

A <surface> element can contain one or more <zone> elements, each of which represents a region or *bounding box* defined in terms of the same coordinate space as that of its parent <surface> element. A zone may be rectangular or non-rectangular: a rectangular zone is defined by a sequence of four coordinates in the same way as a surface; a non-rectangular zone is defined using the attribute *points*, which provides a sequence of coordinates, each of which specifies a point on the perimeter of the zone.[77]

[76] The coordinate space may be thought of as a grid superimposed on a rectangular space. Rectangular areas of the grid are defined as four numbers a b c d: the first two identify the grid point which is at the upper left corner of the rectangle; the second two give the grid point located at the lower right corner of the rectangle. The grid point a b is understood to be the point which is located a points from the origin along the x (horizontal) axis, and b points from the origin along the y (vertical) axis.

[77] The *points* attribute supplies a 'points specification' in the same form as that required by the `<polyline>` or `<polygon>` elements in the SVG standard. See `https://www.w3.org/TR/SVG/shapes.html#PointsBNF`

A zone may be used to define any region of interest, such as a detail or illustration, or some part of the surface which is to be aligned with a particular text element, or otherwise distinguished from the rest of the surface. A surface establishes a coordinate system which may be used to address parts or the whole of some digital representation of a written surface. A zone, by contrast, defines any arbitrary area of interest relative to that surface, using the same coordinate system. It might be bigger or smaller than its parent surface, or might overlap its boundaries. The only constraint is that it must be defined using the same coordinate system.

When an image of some kind is supplied within either a zone or a surface, the implication is that the image represents the whole of the zone or surface concerned. In the simple case therefore, we might imagine a surface defining a page, within which there is a graphic representing the whole of that page, and a number of zones defining parts of the page, each with its own graphic, each representing a part of the page. If however one of those graphics actually represents an area larger than the page (for example to include a binding or the surface of a desk on which the page rests), then it will be enclosed by a zone with coordinates larger than those of the parent surface.

For example, consider the following figure: This is an image of a two page spread from a manuscript in the

Figure 14: Badische Landesbibliothek, Manuscript Durlach 1, Fols. 95v-96r

Badische Landesbibliothek, Karlsruhe. We have no information as to the dimensions of the original object, but the low resolution image displayed here contains 500 pixels horizontally and 321 pixels vertically. For convenience, we might map each pixel to one cell of the coordinate space.[78]

We therefore define a <surface> element corresponding with the area of the image which represents the whole of the two page spread and embed the graphic within it:

```
<facsimile>
  <surface ulx="0" uly="0" lrx="500" lry="321">
    <graphic url="http://upload.wikimedia.org/wikipedia/commons/5/50/Handschrift.karlsruhe.blb.jp
  </surface>
</facsimile>
```

If desired, the <binaryObject> element described in *3.10. Graphics and Other Non-textual Components* (or any other element from the `model.graphicLike` class) may be used instead of a <graphic> element.

Since the image in this example is of a two page opening, we will probably wish to define at least two nested zones, one for each page:

```
<facsimile>
  <surface ulx="50" uly="20" lrx="400"
    lry="280">
    <zone ulx="0" uly="0" lrx="500" lry="321">
      <graphic url="http://upload.wikimedia.org/wikipedia/commons/5/50/Handschrift.karlsruhe.blb.
    </zone>
    <zone ulx="50" uly="20" lrx="210" lry="280">
<!-- left hand page -->
```

[78] The coordinate space used here is based on pixels, but the mapping between pixels and units in the coordinate space need not be one-to-one; it might be convenient to define a more delicate grid, to enable us to address much smaller parts of the image. This can be done simply by supplying appropriate values for the attributes which define the coordinate space; for example doubling them all would map each pixel to two grid points in the coordinate space.

```
    </zone>
    <zone ulx="240" uly="25" lrx="400"
      lry="280">
<!-- right hand page -->
    </zone>
    <zone ulx="90" uly="40" lrx="200" lry="225">
<!--- written part of left hand page -->
    </zone>
  </surface>
</facsimile>
```

As this example shows, in addition to acting as a container for <graphic> elements, <zone> elements may be used to identify parts of a surface for analytical purposes.

The relationship between zone and surface can be quite complex: for example, it may be appropriate to treat the whole of a two page spread as a single written surface, perhaps because particular written zones span both pages. A zone may contain a nested surface, if for example a page has an additional scrap of paper attached to it. A zone may be of any shape, not simply rectangular. Discussion of these and other cases are provided in section *12.2.2.1. Advanced Uses of surface and zone* below.

In the following extended example, we discuss a hypothetical digital edition of an early 16th century French work, Charles de Bovelles' *Géometrie Pratique*.[79] In this edition, each page has been digitized as a separate file: for example, recto page 49 is stored in a file called `Bovelles-49r.png`. In the <facsimile> element used to contain the whole set of pages, we define a <surface> element for this page, which we situate within a coordinate scale running from 0 to 200 in the x (horizontal) axis, and 0 to 300 in the y (vertical) axis. The <surface> element contains a <graphic> element which represents the whole of this surface:

```
<facsimile>
  <surface ulx="0" uly="0" lrx="200" lry="300">
    <graphic url="Bovelles-49r.png"/>
  </surface>
</facsimile>
```

We can now identify distinct zones within the page image using the coordinate scale defined for the surface. In the following figure 15 we show the upper part of the page, with boxes indicating four such zones. Each of these will be represented by a <zone> element, given within the <surface> element already defined, and specified in terms of the same coordinate system. Some zones of interest are indicated by red lines in the following image. The following encoding defines each of the four zones identified in the figure above.

```
<facsimile>
  <surface ulx="0" uly="0" lrx="200" lry="300">
    <graphic url="Bovelles-49r.png"/>
    <zone ulx="25" uly="25" lrx="180" lry="60">
<!-- contains the title -->
    </zone>
    <zone ulx="28" uly="75" lrx="175" lry="178"/>
<!-- contains the paragraph in italics -->
    <zone ulx="105" uly="76" lrx="175"
      lry="160"/>
<!-- contains the figure -->
    <zone ulx="45" uly="125" lrx="60" lry="130"/>
<!-- contains the word "pendans" -->
```

[79] The image is taken from the collection at `https://web.archive.org/web/20191023112411/http://ancilla.unice.fr/illustr.html`, and was digitized from a copy in the Bibliothèque Municipale de Lyon, by whose kind permission it is included here.

De Geometrie. 49

DV SON ET ACCORD DES CLOCHES ET des alleures des cheuaulx, chariotz & charges, des fontaines: & encyclie du monde, & de la dimenſion du corps humain.

Chapitre ſeptieſme.

1 Le ſon & accord des cloches pendans en vng meſme axe, eſt faict en contraires parties.

*Les cloches ont quaſi figures de rondes pyramides imperfaictes & irregulieres: & leur accord ſe fait par reigle geometrique. Comme ſi les deux cloches C & D ſont pendans à vng meſme axe ou eſſieu A B: ie dis que leur accord ſe fera en cõtraires parties, cõme voyez icy figuré. Car quãd lune ſera en hault, laultre declinera embas. Aultrement ſi elles declinent toutes deux enſemble en vne meſme partie, elles feront diſcord, & ſera leur ſonnerie mal plaiſante à oyr.*

C A B D

2 Le vray accord de deux cloches par la touchement des bataulx, eſt faict diametralement.

*Chaſcune cloche a deux diuers coſtez, entre leſquelz le batail touche par diuerſes fois. Et à cauſe que laccord de deux cloches ſe fait par linclination delles en diuerſes & contraires parties: Il fault que ledict accord ſe faſſe diametralemẽt*

ba. dim. cloche. cloche. dim. lan.

G.j.

Figure 15: Detail of p 49r from Bovelles *Géometrie Pratique*

```
  </surface>
</facsimile>
```

Note that the location of each zone is defined independently but using the same coordinate system.

A non rectangular-zone, for example that containing the word *cloche.* at bottom left of the page, could also be defined, using the *points* attribute:

```
<facsimile>
  <surface ulx="0" uly="0" lrx="200" lry="300">
    <graphic url="Bovelles-49r.png"/>
    <zone points="4.8,31.0 5.4,30.7 5.5,32.2 5.8,32.8 6.1,33.4 5.5,33.7 5.1,33.3
4.6,32.2"/>
  </surface>
</facsimile>
```

In the examples above a single <graphic> element has been associated directly with the surface of the page rather than nesting it within a zone. However, it is also possible to include multiple <zone> elements which contain a <graphic> element, if for example a detailed image is available. Since all <zone> elements use the same coordinate system (that defined by their parent <surface>), there is no need to demonstrate enclosure of one zone within another by means of nesting. To continue the current example, supposing that we have an additional image called `Bovelles49r-detail.png` containing an additional image of the figure in the third zone above, we might encode that zone as follows:

```
<zone ulx="105" uly="76" lrx="175" lry="160">
  <graphic url="Bovelles49r-detail.png"/>
</zone>
```

Within a <surface> or <zone>, individual lines can be identified using the <path> element, which also carries the *points* attribute:

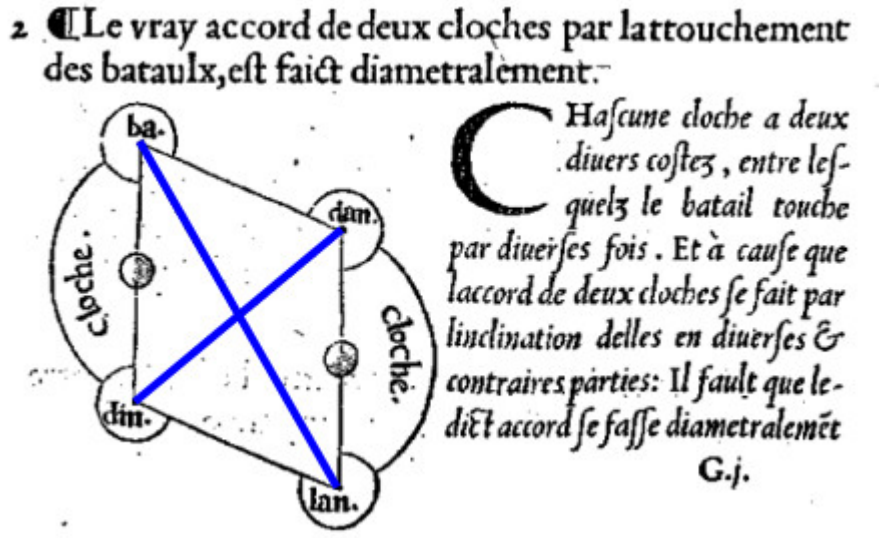



Figure 16: Smaller detail of p 49r from Bovelles highlighting two specific lines

```
<surface ulx="0" uly="0" lrx="443" lry="272">
  <graphic url="facs-fig3.jpg"/>
  <path points="74,73 171,244"
    xml:id="balan"/>
  <path points="71,203 173,116"
    xml:id="dindan"/>
</surface>
```

This is useful for linking an annotation or explanation to a specific line on an object surface. Any number of coordinates can be included to specify lines which are not straight; this example shows how the first of the famous 'story lines' appearing at the beginning of chapter 40 of *Tristram Shandy* might be encoded:

```
<surface ulx="0" uly="0" lrx="411" lry="652">
  <graphic url="sterne.png"/>
  <path points="65,511 88,510 92,517 100,521 107,520 112,516 116,512 117,511 152,512
156,508 162,505 169,505 174,506 178,509 179,512 236,514 136,479 208,493 270,512
328,513 336,525 339,528 339,536 331,539 328,535 329,530 334,526 342,521 345,519
350,512 397,514 402,498 414,515 425,515 435,531 440,513 475,513 475,518 477,520
479,521 481,522 483,521 484,518 486,516 486,514 491,512 494,514 496,520 496,529
493,535 494,539 497,543 501,543 504,543 507,540 508,537 507,526 505,518 502,510
501,508 501,503 503,501 506,500 510,500 512,503 513,507 511,513 543,516 552,513
552,501 550,496 549,490 552,486 562,487 564,468 559,465 557,462 556,457 558,453
562,450 570,451 573,446 579,433"/>
</surface>
```

## 12.2 Combining Transcription with Facsimile

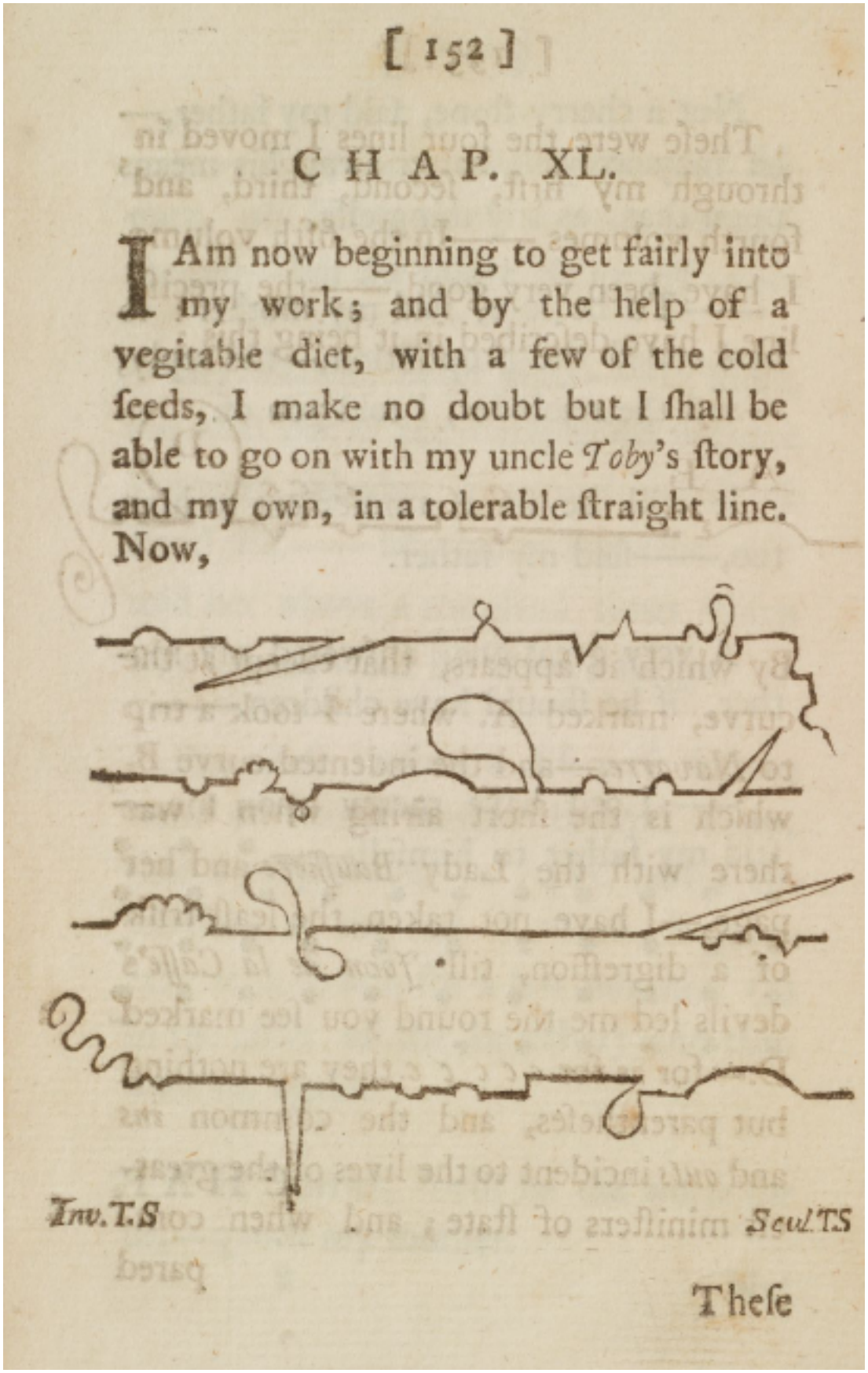
A digitized source document may contain nothing more than page images and a small amount of metadata. It may also contain an encoded transcription of the pages represented, which may either be 'embedded' within a <sourceDoc> element, or supplied in parallel with a <facsimile> as defined above.

If the transcription is regarded as a text in its own right, organized and structured independently of its physical realization in the document or documents represented by the facsimile, then the recommended practice is to use the <text> element to contain such a structured representation, and to present it in parallel. The <text> element is a sibling of the <facsimile> and <sourceDoc> elements. This approach is illustrated in section *12.2.1. Parallel Transcription* below. Alternatively, if the transcription is intended not to prioritize representation of the final text so much as the process by which the document came to take its present form, or the physical disposition of its component parts, it can be presented as an embedded transcription, as further described in section *12.2.2. Embedded Transcription* below.

### 12.2.1 Parallel Transcription

Suppose now that we wish to align a transcription of the page discussed in the preceding section with particular zones. We begin by giving each relevant part of the facsimile an identifier:

```
<facsimile>
  <surface ulx="0" uly="0" lrx="200" lry="300">
    <zone xml:id="B49r" ulx="0" uly="0"
      lrx="200" lry="300">
      <graphic url="Bovelles-49r.png"/>
    </zone>
    <zone ulx="105" uly="76" lrx="175"
      lry="160">
      <graphic url="Bovelles49r-detail.png"/>
    </zone>
```



[ 152 ]

CHAP. XL.

I Am now beginning to get fairly into my work; and by the help of a vegitable diet, with a few of the cold ſeeds, I make no doubt but I ſhall be able to go on with my uncle *Toby*'s ſtory, and my own, in a tolerable ſtraight line. Now,

*Inv. T. S* *Scul. T. S*

Theſe

Figure 17: Page 152 of *Tristram Shandy*

```
   <zone xml:id="B49rHead" ulx="25" uly="25"
     lrx="180" lry="60"/>
<!-- contains the title -->
   <zone xml:id="B49rPara2" ulx="28" uly="75"
     lrx="175" lry="178"/>
<!-- contains the first paragraph in italics -->
   <zone xml:id="B49rFig1" ulx="105" uly="76"
     lrx="175" lry="160"/>
<!-- contains the figure -->
   <zone xml:id="B49rW457" ulx="45" uly="125"
     lrx="60" lry="130"/>
<!-- contains the word "pendans" -->
  </surface>
</facsimile>
```

The alignment between transcription and image is made, as usual, by means of the *facs* attribute:

```
<pb facs="#B49r"/>
<fw>De Geometrie 49</fw>
<head facs="#B49rHead"> DU SON ET ACCORD DES CLOCHES ET <lb/> des alleures des
chevaulx, chariotz & charges, des fontaines:& <lb/> encyclie du monde,
& de la dimension du corps humain.<lb/> Chapitre septiesme</head>
<div n="1">
  <p>Le son & accord des cloches pendans en ung mesme <lb/> axe, est faict en
     contraires parties.</p>
  <p rend="it" facs="#B49rPara2">LEs cloches ont quasi fi<lb/>gures de rondes
     pyra<lb/>mides imperfaictes & <lb/> irregulieres: & leur accord se
  <lb/> fait par reigle geometrique. Com<lb/>me si les deux cloches C & D
  <lb/> sont <w facs="#B49rW457">pendans</w> à ung mesme axe <lb/> ou essieu A B:
     je dis que leur ac<lb/>cord se fera en co<ex>n</ex>traires parties<lb/>
     co<ex>m</ex>me voyez icy figuré. Car qua<ex>n</ex>d <lb/> lune sera en
     hault, laultre declinera embas. Aultrement si elles decli<lb/>nent toutes deux
     ensembles en une mesme partie, elles seront discord, <lb/> & sera leur
     sonnerie mal plaisante à oyr.<figure facs="#B49rFig1">
      <graphic url="Bovelles49r-detail.png"/>
    </figure>
  </p>
</div>
```

It is also possible to point in the other direction, from a <surface> or <zone> to the corresponding text. This is the function of the *start* attribute, which supplies the identifier of the element containing at least the start of the transcribed text found within the surface or zone concerned. Thus, another way of linking this page with its transcription would be simply

```
<facsimile>
  <surface start="#PB49R">
    <graphic url="Bovelles-49r.png"/>
  </surface>
</facsimile>
<text>
  <body>
    <div>
<!-- ... -->
     <pb xml:id="PB49R"/>
```

```
     <fw>De Geometrie 49</fw>
<!-- ... -->
   </div>
  </body>
</text>
```

### 12.2.2 Embedded Transcription

An *embedded transcription* is one in which words and other written traces are encoded as subcomponents of elements representing the physical surfaces carrying them rather than independently of them.

The following elements are available for this purpose:

`<sourceDoc>` contains a transcription or other representation of a single source document potentially forming part of a *dossier génétique* or collection of sources.

`<surface>` defines a written surface as a two-dimensional coordinate space, optionally grouping one or more graphic representations of that space, zones of interest within that space, and, when using an embedded transcription approach, transcriptions of the writing within them.

`<zone>` defines any two-dimensional area within a <surface> element.

`<line>` contains the transcription of a topographic line in the source document.

`<seg>` (arbitrary segment) represents any segmentation of text below the 'chunk' level.

The elements <surface>, <surfaceGrp>, and <zone> were introduced above in section *12.1. Digital Facsimiles*. When supplied within a <sourceDoc> element, these elements may contain transcriptions of the written content of a source in addition to or as an alternative to digital images of them. Such transcription may be placed directly within the <zone> element, or within one or more <line> elements, for cases where the writing is linear, in the sense that it is composed of discrete tokens organized physically into groups, typically organized in a sequence corresponding with the way they are intended to be read. Depending on the directionality of the writing system used, this might be any combination of top-down and left to right, or vice versa. The element <line> may be used to hold a complete group of such tokens. Where, however, the lineation is not considered significant, any group of tokens may be indicated using the <zone> element. The <seg> element described in section *17.3. Blocks, Segments, and Anchors* may also be used to indicate smaller sequences of tokens within <zone>, or <line> as appropriate.

Returning to the preceding example, we might transcribe the content of the zone to which we gave the identifier `B49rPara2` within a <sourceDoc> element as follows:

```
<sourceDoc>
  <surface ulx="0" uly="0" lrx="200" lry="300">
    <zone ulx="0" uly="0" lrx="200" lry="300">
      <graphic url="Bovelles-49r.png"/>
    </zone>
<!-- ... -->
    <zone ulx="28" uly="75" lrx="175" lry="178">
      <line>LEs cloches ont quasi
            fi</line>
      <line>gures de rondes pyra</line>
      <line>mides imperfaictes &
      </line>
      <line> irregulieres: & leur accord se</line>
      <line> fait par reigle geometrique. Com</line>
      <line>me si les deux cloches C
            & D </line>
      <line> sont <zone ulx="45" uly="125" lrx="60"
         lry="130">pendans</zone> à ung mesme axe</line>
      <line> ou essieu A B: je dis que
            leur ac</line>
      <line>cord se fera en cõtraires parties</line>
      <line> cõme
```

```
          voyez icy figuré. Car quãd </line>
     <line> lune sera en hault, laultre
          declinera embas. Aultrement si elles declinent toutes deux ensembles en une
          mesme partie, elles seront discord,</line>
     <line> & sera leur sonnerie
          mal plaisante à oyr.</line>
   </zone>
   <zone ulx="105" uly="76" lrx="175"
     lry="160">
     <graphic url="Bovelles49r-detail.png"/>
   </zone>
  </surface>
</sourceDoc>
```

As mentioned above, some or all of the written surfaces being transcribed may be composed of physically distinct scraps. In the following example, taken from the Walt Whitman Archive, two pieces of newsprint have been glued to a piece of blue paper on which a poem is being drafted: The two pieces of newsprint might simply be regarded as special kinds of zone, but they are also new surfaces, since they might contain additional written zones themselves (such as the numbers in this case).

Using these elements, the Whitman draft above might be encoded as follows:

```
<surface>
  <zone>
    <line>Poem</line>
    <line>As in Visions of – at</line>
    <line>night –</line>
    <line>All sorts of fancies running through</line>
    <line>the head</line>
  </zone>
  <zone>
    <surface type="newsprint"
      attachment="glue" flipping="false">
      <zone>Spring has just set in here, and the weather[...] a steamer </zone>
      <metamark function="sequence">2</metamark>
    </surface>
  </zone>
  <zone>
    <surface type="newsprint"
      attachment="glue" flipping="false">
      <zone>"The shores on either side of the Sound are... The In- </zone>
      <metamark function="sequence">3</metamark>
    </surface>
  </zone>
</surface>
```

The <metamark> element used in this example is further discussed below (*12.3.4.2. Metamarks*)

Note that in this example we have not included any <graphic> element corresponding with the <zone> or <surface> elements identified in the transcription. The encoder may choose to complement a transcription with graphic representations of its source at whatever level is considered effective, or not at all. Equally, the encoder may choose to provide only graphics without any transcription, to provide only a structured (non-embedded) transcription, or to provide any combination of the three.

This example also lacks any coordinate information to specify either the size of the two newspaper fragments or whereabouts on the parent <surface> element they are to be found, other than the reading order implicit in their sequence. Such information could be added if desired by specifying a coordinate system on the outermost <surface>

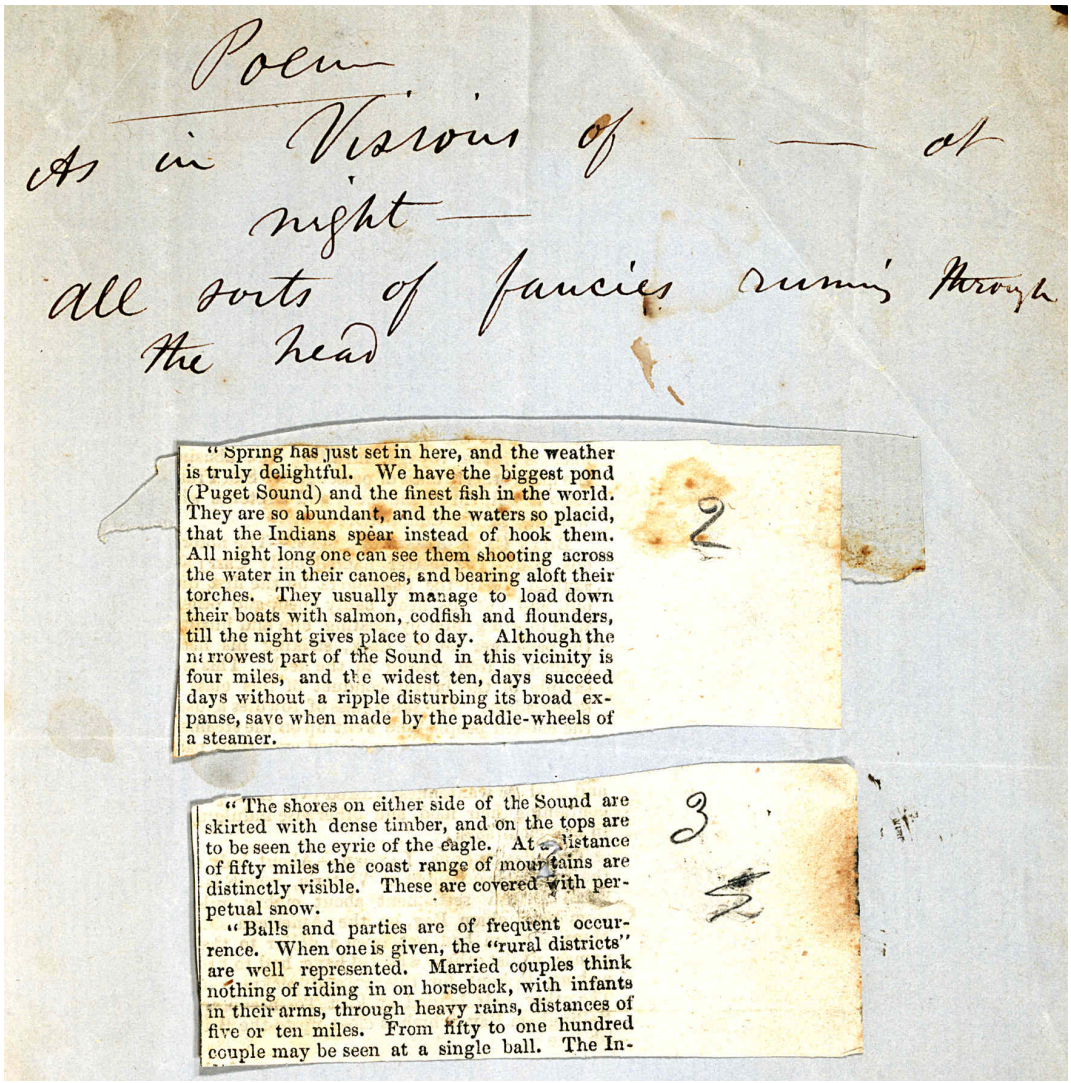



Figure 18: Single leaf of notes possibly related to the poem eventually titled Sleepers. From the Walt Whitman Archive (Duke 258).

element, and then indicating values within that system for each of the two fragments, as was discussed above. We discuss this in further detail in section *12.2.2.1. Advanced Uses of surface and zone* below.

#### 12.2.2.1 Advanced Uses of <surface> and <zone>

As a child of <sourceDoc>, the <surface> element both identifies a specific area containing writing and provides a two dimensional set of coordinates which can be used to position and define dimensions for sub-parts of it. Furthermore, surfaces may nest within other surfaces, as in the case of 'patches' or other written materials attached to the main writing surface. In the general case, the position and dimensions of such nested surfaces will be defined using the same coordinate system as that supplied by the parent <surface> element. It is also possible, however, that a different coordinate system is required for such a nested surface, perhaps because it requires a more complex granularity. We consider both possibilities.

In the earlier examples showing nested examples we did not provide any coordinate information, for simplicity of presentation. Suppose however, that we wish to indicate the position and sizes of the newspaper scraps in 18 above, relative to the whole page. As previously noted, the four attributes *ulx*, *uly*, *lrx* and *lry* when given on the <surface> element define the coordinate scheme, rather than specifying the location of that surface. We must therefore introduce an additional <zone> element, as in the following revised encoding for this example

```
<surface ulx="0" uly="0" lrx="50" lry="50">
  <zone ulx="1" uly="1" lrx="10" lry="10">
    <line>Poem</line>
    <line>As in Visions of – at</line>
    <line>night –</line>
    <line>All sorts of fancies running through</line>
    <line>the head</line>
  </zone>
  <zone ulx="4" uly="4" lrx="20" lry="20">
    <surface type="newsprint"
      attachment="glue" flipping="false">
      <zone>Spring has just set in here, and the weather[...] a steamer </zone>
      <metamark function="sequence">2</metamark>
    </surface>
  </zone>
</surface>
```

In this version of the encoding, the inner surface, corresponding with the first piece of newsprint, inherits locational information from the <zone> element that contains it. This zone, and the preceding one, which contains a sequence of <line> elements, are both positioned in terms of the coordinates specified on the outermost <surface> element, which defines a scale running from 0 to 50 in either direction. On that scale, the <line> elements occupy a rectangle with coordinates (1,1,10,10), while the nested surface occupies a rectangle with coordinates (4,4,20,20).

Now suppose that we wish to define a finer scale grid for the newspaper patch, perhaps because we wish to localize zones within it with greater accuracy. To do this we will need to specify the position of the nested surface as in the previous example, but also to define the new coordinate system. We accomplish this as follows:

```
<surface ulx="0" uly="0" lrx="50" lry="50">
  <zone ulx="1" uly="1" lrx="10" lry="10">
    <line>Poem</line>
<!-- ... -->
    <line>the head</line>
  </zone>
  <zone ulx="4" uly="4" lrx="20" lry="20">
    <surface ulx="0" uly="0" lrx="100"
      lry="100">
      <zone ulx="10" uly="10" lrx="90" lry="95"> Spring has just set in here, and the
          weather [...] a steamer </zone>
    </surface>
```

```
  </zone>
</surface>
```

As before, the second zone defines the position and size of the newspaper patch itself in terms of a coordinate system running from 0 to 50 on both X and Y axes. The nested <surface> element however defines a new scale for all of its components, running from 0 to 100 on both X and Y axes. The position of the nested zone containing the text *Spring … steamer* is now given in terms of this scale.

All of the examples so far given have involved rectangular zones, for clarity of exposition. As noted above, the *points* attribute may be used to define non-rectangular zones as a series of points. For example, in the last of the Whitman examples discussed in section *12.3.4.2. Metamarks* above, we might wish to record the exact shape of the zone containing the metamark *Entered*. Since this is not a rectangular zone, we use the *points* attribute to indicate the points defining a polygon which contains it. The values used are expressed in terms of a coordinate space running from 0 to 229 in the X dimension, and 0 to 160 in the Y dimension.

```
<surface ulx="0" uly="0" lrx="229" lry="160">
  <graphic url="whitman-02.jpg"/>
  <zone xml:id="entered"
    points="142,122 155,113 178,122 208,144 198,154 178,139"/>
</surface>
```

In exactly the same way, we may wish to identify the curved zone in the following image containing the word *Northamptonshire*: This curved zone might be encoded in the following way:

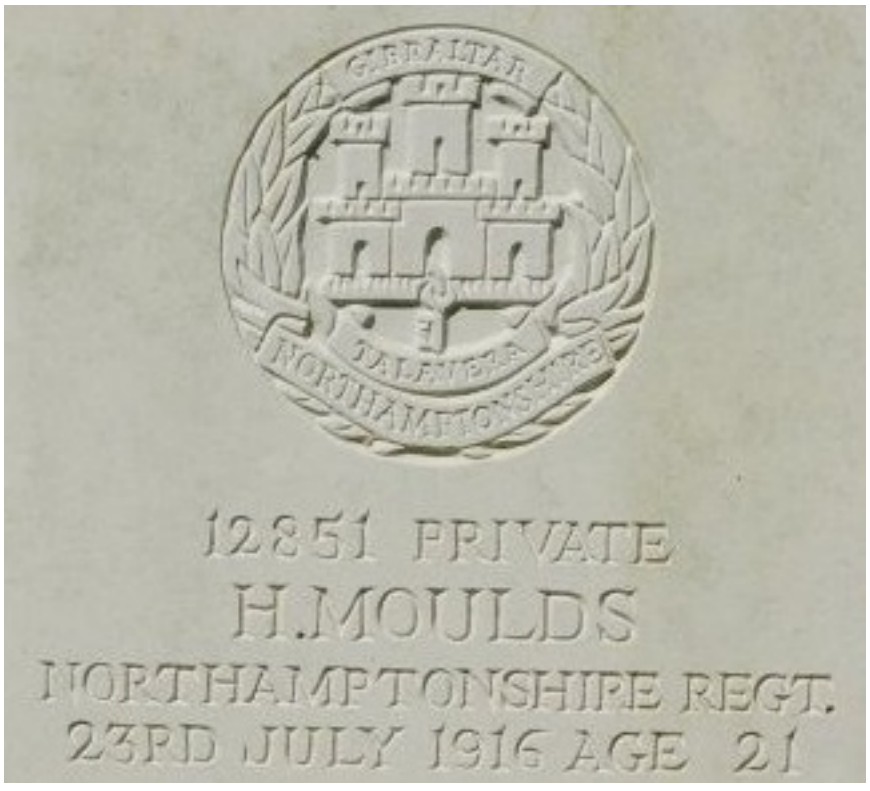



Figure 19: Gravestone of Private Moulds

```
<surface xml:id="badge" ulx="14.54"
  uly="16.14" lrx="0" lry="0">
  <graphic url="stone.jpg"/>
  <zone xml:id="county"
    points="4.6,6.3 5.25,5.85 6.2,6.6 8.2,7.4 9.9,6.6 10.9,6.1 11.4,6.7 8.2,8.3
6.2,7.6"/>
</surface>
```

Finally, it should be noted that a <zone> does not need to be entirely contained within the two-dimensional space defined by its parent surface. For example, we might wish to encode the example in 14 above not as a surface representing the whole of the two page spread, but as a surface representing only the written part of this opening. The written part appears 50 units from the left of the image and 20 units from the top, while the bottom right corner of the written part appears 400 units from the left of the image, and 280 units from the top. We therefore define the written surface within this image as follows:

```
<facsimile>
  <surface ulx="50" uly="20" lrx="400"
    lry="280">
<!-- ... -->
  </surface>
</facsimile>
```

To describe the whole image, we will now need to define a zone of interest which represents an area larger than this surface. Using the same coordinate system as that defined for the surface, its coordinates are 0,0,500,321. This zone of interest can be defined by a <zone> element, within which we can place the uncropped <graphic>:

```
<facsimile>
  <surface ulx="50" uly="20" lrx="400"
    lry="280">
    <zone ulx="0" uly="0" lrx="500" lry="321">
      <graphic url="http://upload.wikimedia.org/wikipedia/commons/5/50/Handschrift.karlsruhe.blb.
    </zone>
  </surface>
</facsimile>
```

## 12.3 Scope of Transcriptions

When transcribing a primary source, whether using <text> or <sourceDoc>, scholars may wish to record information concerning individual readings of letters, words, or larger units.They may also wish to include other editorial material, such as comments on the status or possible origin of particular readings, corrections, or text supplied to fill lacunae.

Such elements may also be used for digital transcriptions in which the object is not to represent a finished text, but rather to represent the creative process, as evidenced by different 'layers' or 'traces' of writing in one or more documents. Transcriptions of this kind are closely focussed on the physical appearance of specific documents, needing to distinguish the traces of different writing activities on them, such as additions and deletions but also other indications of how the writing is to be read, such as indications of transposition, re-affirmation of writing which has been deleted, and so on. Such distinctions are considered of particular importance when dealing with authorial manuscripts, but are also relevant in the case of historical sources such as charters or other legal documents.

In either case, it is customary in transcriptions to register certain features of the source, such as ornamentation, underlining, deletion, areas of damage and lacunae. This chapter provides ways of encoding such information:

- methods of recording editorial or other alterations to the text, such as expansion of abbreviations, corrections, conjectures, etc. (section *12.3.1. Altered, Corrected, and Erroneous Texts*)

- methods of describing important extra-linguistic phenomena in the source: pagination, lineation, unusual spaces, changes of manuscript hand, etc. (section *12.3.2. Hands and Responsibility*)
- methods of representing aspects of layout such as spacing or lines *12.4. Aspects of Layout*
- methods of representing material such as running heads, catch-words, and the like (section *12.6. Headers, Footers, and Similar Matter*)

The remainder of this chapter describes a model for encoding such transcriptions, in which elements such as <mod>, <del>, etc. are used to mark writing traces and their functions within the document. Each such element can be assigned to one or more editorially-defined modification groups, termed a *change*, by means of a global *change* attribute, which references a definition for the modification group concerned, typically provided within the TEI header <creation> element; see further *12.7. Identifying Changes and Revisions*. The transcription itself may be embedded within the elements <surface> and <zone> described in section *12.1. Digital Facsimiles*, or provided in parallel within a <text> element. Within a <zone>, the transcription may be organized topographically in terms of lines of writing, using the <line> element, or in terms of further nested zones, or as a combination of the two (*12.2.2. Embedded Transcription*).

These recommendations are not intended to meet every transcriptional circumstance likely to be faced by any scholar. Rather, they should be regarded as a base which can be elaborated if necessary by different scholars in different disciplines

As a rule, all elements which may be used in the course of a transcription of a single witness may also be used in a critical apparatus, i.e. within the elements proposed in chapter *13. Critical Apparatus*. This can generally be achieved by nesting a particular reading containing tagged elements from a particular witness within the <rdg> element in an <app> structure.

Just as a critical apparatus may contain transcriptional elements within its record of variant readings in various witnesses, one may record variant readings in an individual witness by use of the apparatus mechanisms <app> and <rdg>. This is discussed in section *13.3. Using Apparatus Elements in Transcriptions*.

## 12.3.1 Altered, Corrected, and Erroneous Texts

In the detailed transcription of any source, it may prove necessary to record various types of actual or potential alteration of the text: expansion of abbreviations, correction of the text (either by author, scribe, or later hand, or by previous or current editors or scholars), addition, deletion, or substitution of material, and similar matters. The sections below describe how such phenomena may be encoded using either elements defined in the core module (defined in chapter *3. Elements Available in All TEI Documents*) or specialized elements available only when the module described in this chapter is available.

### 12.3.1.1 Core Elements for Transcriptional Work

In transcribing individual sources of any type, encoders may record corrections, normalizations, additions, and omissions using the elements described in section *3.5. Simple Editorial Changes*. Representation of abbreviations and their expansions may also involve use of elements described in section *3.6. Names, Numbers, Dates, Abbreviations, and Addresses*. Elements particularly relevant to this chapter include:

`<abbr>` (abbreviation) contains an abbreviation of any sort.

`<add>` (addition) contains letters, words, or phrases inserted in the source text by an author, scribe, or a previous annotator or corrector.

`<choice>` (choice) groups a number of alternative encodings for the same point in a text.

`<corr>` (correction) contains the correct form of a passage apparently erroneous in the copy text.

`<del>` (deletion) contains a letter, word, or passage deleted, marked as deleted, or otherwise indicated as superfluous or spurious in the copy text by an author, scribe, or a previous annotator or corrector.

`<expan>` (expansion) contains the expansion of an abbreviation.

`<gap>` (gap) indicates a point where material has been omitted in a transcription, whether for editorial reasons described in the TEI header, as part of sampling practice, or because the material is illegible, invisible, or inaudible.

`<sic>` (Latin for thus or so) contains text reproduced although apparently incorrect or inaccurate.

All of these elements bear additional attributes for specifying who is responsible for the interpretation

represented by the markup, and the associated certainty. In addition, some of them bear an attribute allowing the markup to be categorized by type and source.

`att.editLike` provides attributes describing the nature of an encoded scholarly intervention or interpretation of any kind.

- `@evidence` indicates the nature of the evidence supporting the reliability or accuracy of the intervention or interpretation.

`att.global.source` provides attributes used by elements to point to an external source.

- `@source` specifies the source from which some aspect of this element is drawn.

`att.global.responsibility` provides attributes indicating the agent responsible for some aspect of the text, the markup or something asserted by the markup, and the degree of certainty associated with it.

- `@cert` (certainty) signifies the degree of certainty associated with the intervention or interpretation.
- `@resp` (responsible party) indicates the agency responsible for the intervention or interpretation, for example an editor or transcriber.

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

- `@type` characterizes the element in some sense, using any convenient classification scheme or typology.
- `@subtype` (subtype) provides a sub-categorization of the element, if needed.

The specific aspect of the markup described by these attributes differs on different elements; for further discussion, see the relevant sections below, especially section *12.3.2.2. Hand, Responsibility, and Certainty Attributes*.

The following sections describe how the core elements just named may be used in the transcription of primary source materials.

#### 12.3.1.2 Abbreviation and Expansion

The writing of manuscripts by hand lends itself to the use of abbreviation to shorten scribal labour. Commonly occurring letters, groups of letters, words, or even whole phrases, may be represented by significant marks. This phenomenon of manuscript abbreviation is so widespread and so various that no taxonomy of it is here attempted. Instead, methods are shown which allow abbreviations to be encoded using the core elements mentioned above.

A manuscript abbreviation may be viewed in two ways. One may transcribe it as a particular sequence of letters or marks upon the page: thus, a 'p with a bar through the descender', a 'superscript hook', a 'macron'. One may also interpret the abbreviation in terms of the letter or letters it is seen as standing for: thus, 'per', 're', 'n'. Both of these views are supported by these Guidelines.

In many cases the glyph found in the manuscript source also exists in the Unicode character set: for example the common Latin brevigraph ꝛ, standing for *et* and often known as the 'Tironian et' can be directly represented in any XML document as the Unicode character with code point U+204A (see further *v.7.1 Character References* and *vi.1 Language Identification*). In cases where it does not, these Guidelines recommend use of the <g> element provided by the `gaiji` module described in chapter *5. Characters, Glyphs, and Writing Modes*. This module allows the encoder great flexibility both in processing and in documenting non-standard characters or glyphs, including the ability to provide detailed documentation and images for them.

These two methods of coding abbreviation may also be combined. An encoder may record, for any abbreviation, both the sequence of letters or marks which constitutes it, and its sense, that is, the letter or letters for which it is believed to stand. For example, in the following fragment the phrase *euery persone* is represented by a sequence of abbreviated characters: These lines may be transcribed directly, using the <g> element to indicate the two brevigraphs as follows:

```
eu<g  ref="#b-er">er</g>y <g  ref="#b-per">per</g>sone that loketh after heuen hath a
place in
this ladder
<!-- elsewhere -->
<charDecl>
  <char  xml:id="b-er">
<!-- definition for the er brevigraph -->
```

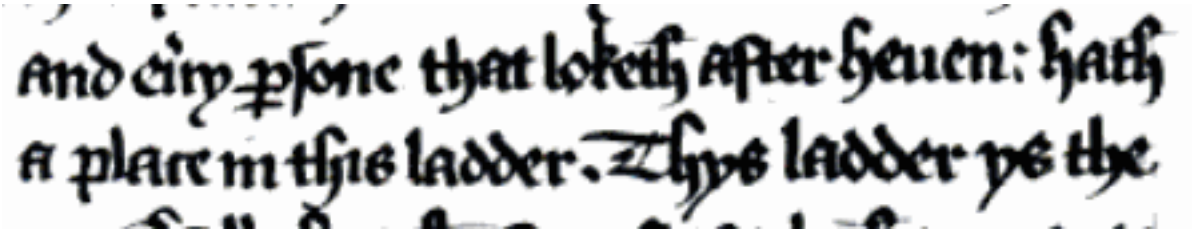



Figure 20: Detail from fol. 126v of Bodleian MS Laud Misc 517

```
  </char>
  <char  xml:id="b-per">
<!-- definition for the per brevigraph -->
  </char>
</charDecl>
```

Note that in each case the <g> element may contain a suggested replacement for the referenced brevigraph; this is purely advisory however, and may not be appropriate in all cases. The referenced character definitions may be located elsewhere in this or some other document, typically forming part of a <charDecl> element, as described in *5.2. Markup Constructs for Representation of Characters and Glyphs*.

The transcriber may also wish to indicate that, because of the presence of these particular characters, the two words are actually abbreviations, by using the <abbr> element:

```
<abbr>eu<g  ref="#b-er">er</g>y</abbr>
<abbr>
  <g  ref="#b-per">per</g>sone
</abbr> ...
```

Alternatively, the transcriber may choose silently to expand these abbreviations, using the <expan> element:

```
<expan>euery</expan>
<expan>persone</expan> ...
```

And, of course, the <choice> element can be used to show that one encoding is an alternative for the other:

```
<choice>
  <abbr>eu<g  ref="#b-er">er</g>y</abbr>
  <expan>euery</expan>
</choice>
```

When abbreviated forms such as these are expanded, two processes are carried out: some characters not present in the abbreviation are added (always), and some characters or glyphs present in the abbreviation are omitted or replaced (often). For example, when the abbreviation *Dr.* is expanded to *Doctor*, the dot in the abbreviation is removed, and the letters *octo* are added. Where detailed markup of abbreviated words is required, these two aspects may be marked up explicitly, using the following elements:

<ex> (editorial expansion) contains a sequence of letters added by an editor or transcriber when expanding an abbreviation.

<am> (abbreviation marker) contains a sequence of letters or signs present in an abbreviation which are omitted or replaced in the expanded form of the abbreviation.

Using these elements, a transcriber may indicate the status of the individual letters or signs within both the abbreviation and the expansion. The <am> element surrounds characters or signs such as tittles or tildes, used to indicate the presence of an abbreviation, which are typically removed or replaced by other characters in the expanded form of the abbreviation:

```
<abbr>eu<am>
    <g  ref="#b-er"/>
  </am>y</abbr>
<abbr>
  <am>
    <g  ref="#b-per"/>
  </am>sone
</abbr> ...
```

while the <ex> element may be used to indicate those characters within the expansion which are not present in the abbreviated form.

```
<expan>eu<ex>er</ex>y</expan>
<expan>
  <ex>per</ex>sone
</expan> ...
```

The content of the <abbr> element should usually include the whole of the abbreviated word, while the <expan> element should include the whole of its expansion. If this is not considered necessary, the <am> and <ex> elements may be used within a <choice> element, as in this example:

```
eu<choice>
  <am>
    <g  ref="#b-er"/>
  </am>
  <ex>er</ex>
</choice>y <choice>
  <am>
    <g  ref="#b-per"/>
  </am>
  <ex>per</ex>
</choice>sone ...
```

As implied in the preceding discussion, making decisions about which of these various methods of representing abbreviation to use will form an important part of an encoder's practice. As a rule, the <abbr> and <am> elements should be preferred where it is wished to signify that the content of the element is an abbreviation, without necessarily indicating what the abbreviation may stand for. The <ex> and <expan> elements should be used where it is wished to signify that the content of the element is not present in the source but has been supplied by the transcriber, without necessarily indicating the abbreviation used in the original. The decision as to which course of action is appropriate may vary from abbreviation to abbreviation; there is no requirement that the same system be used throughout a transcription, although doing so will generally simplify processing. The choice is likely to be a matter of editorial policy. If the highest priority is to transcribe the text *literatim* (letter by letter), while indicating the presence of abbreviations, the choice will be to use <abbr> or <am> throughout. If the highest priority is to present a reading transcription, while indicating that some letters or words are not actually present in the original, the choice will be to use <ex> or <expan> throughout.

Further information may be attached to instances of these elements by the <note> element, on which see section *3.9. Notes, Annotation, and Indexing*, and by use of the *resp* and *cert* attributes. In this instance from the English *Brut*, a note is attached to an editorial expansion of the tail on the final d of *good* to *goode*:

```
For alle the while
that I had good<ex  xml:id="exp01">e</ex> I was welbeloued
```

Then the note:

```
<note  target="#exp01">The stroke added to
the final d could signify the plural ending (-es, -is, -ys>) but the
singular <hi  rend="it">goode</hi> was used with the meaning <q>property</q>,
<q>wealth</q>, at this time (v. examples quoted in OED, sb. Good, C. 7, b,
c, d and 8 spec.)</note>
```

The editor might declare a degree of certainty for this expansion, based on the OED examples, and state the responsibility for the expansion:

```
For alle the while that I
had good<ex  resp="#mp"  cert="high">e</ex> I was welbeloued
```

The value supplied for the *resp* attribute should point to the name of the editor responsible for this and possibly other interventions; an appropriate element therefore might be a <respStmt> element in the header like the following:

```
<respStmt  xml:id="mp">
  <resp>Editorial emendations</resp>
  <name>Malcom Parkes</name>
</respStmt>
```

Observe that the *cert* and *resp* attributes are used with the <ex> element only to indicate confidence in the content of the element (i.e. the expansion), and responsibility for suggesting this expansion respectively.

The <choice> element may be used to indicate that the proposed expansion is one way of encoding what might equally well be represented as an abbreviation, represented by the hooked D, as follows:

```
For alle the while that I had
<choice>
  <sic>goo<abbr>ɗ</abbr>
  </sic>
  <expan  resp="#mp"  cert="high">good<ex>e</ex>
  </expan>
</choice> I was
welbeloued
```

If it is desired to express aspects of certainty and responsibility for some other aspect of the use of these elements, then the mechanisms discussed in chapter *22. Certainty, Precision, and Responsibility* should be used. See also *12.3.2.2. Hand, Responsibility, and Certainty Attributes* for discussion of the issues of certainty and responsibility in the context of transcription.

If more than one expansion for the same abbreviation is to be recorded, multiple notes may be supplied. It may also be appropriate to use the markup for critical apparatus; an example is given in section *13.3. Using Apparatus Elements in Transcriptions*.

#### 12.3.1.3 Correction and Conjecture

The <sic>, <corr>, and <choice> elements, defined in the `core` module should be used to indicate passages deemed in need of correction, or actually corrected, during the transcription of a source. For example, in the manuscript of William James's *A Pluralistic Universe* as edited by Fredson Bowers (Cambridge: Harvard University Press, 1977), a sentence first written

> One must have lived longer with this system, to appreciate its advantages.

has been modified by James to begin 'But One must …', without the initial capital O having been reduced to lowercase. This non-standard orthography could be recorded thus:

```
But <sic>One</sic> must have lived
...
```

or corrected:

```
But
<corr>one</corr> must have lived ...
```

or the two possibilities might be represented as a choice:

```
But
<choice>
  <sic>One</sic>
  <corr>one</corr>
</choice> must have lived
...
```

Similarly, in this example from Albertus Magnus, both a manuscript error *angues* and its correction *augens* are registered within a <choice> element:

```
Nos autem iam
ostendimus quod nutrimentum et
<choice>
  <sic>angues</sic>
  <corr>augens</corr>
</choice>.
```

Note that the <corr> element is used to provide a corrected form which is *not* present in the source; in the case of a correction made in the source itself, whether scribal, authorial, or by some other hand, the <add>, <del>, and <subst> elements described in *12.3.1.4. Additions and Deletions* should be used.

The <sic> element is used to mark passages considered by the transcriber to be erroneous; in such cases, the <corr> element indicates the transcriber's correction of them. Where the transcriber considers that one or more words have been erroneously omitted in the original source and corrects this omission, the <supplied> element discussed in *12.3.1.7. Text Omitted from or Supplied in the Transcription* should be used in preference to <corr>. Thus, in the following example, from George Moore's draft of additional materials for *Memoirs of My Dead Life*, the transcriber supplies the word *we* omitted by the author:

```
You see that I avoid
the word create for we create nothing <supplied>we</supplied> develope.
```

As with <expan> and <abbr>, the choice as to whether to record simply that there is an apparent error, or simply that a correction has been applied, or to record both possible readings within a <choice> element is left to the encoder. The decision is likely to be a matter of editorial policy, which might be applied consistently throughout or decided case by case. If the highest priority is to present an uncorrected transcription while noting perceived errors in the original, the choice will typically be to use only <sic> throughout. If the highest priority is to present a reading

transcription, while indicating that perceived errors in the original have been corrected, the choice will be to use only <corr> throughout.

Further information may be attached to instances of these elements by the <note> element and *resp* and *cert* attributes. Instances of these elements may also be classified according to any convenient typology using the *type* attribute.

For example, consider the following encoding of an emendation in the Hengwrt manuscript proposed by E. Talbot Donaldson:

```
Telle me also, to what conclusioun Were
membres maad, of generacioun And of so parfit wis a <choice  xml:id="corr117">
  <sic>wight</sic>
  <corr>wright</corr>
</choice> ywroght?
<!-- ... -->
<note  target="#corr117">This emendation of the Hengwrt copy text, based on a Latin
source and on the reading of three late and usually unauthoritative
manuscripts, was proposed by E. Talbot Donaldson in
<bibl>
    <title>Speculum</title> 40 (1965) 626–33.</bibl>
</note>
```

The <note> element discussed in *3.9. Notes, Annotation, and Indexing* may be used to give a more detailed discussion of the motivation for or scope of a correction. If linked by means of a pointer (as in this example) it may be located anywhere convenient within the transcription; typically all detailed notes will be collected together in a separate  element in the <back>. Alternatively, the pointer may be omitted, and the <note> placed immediately adjacent to the element being annotated. The advantage of the former solution is that it permits the same annotation to refer to several corrections, by supplying more than one pointer in the *target* attribute of the <note>, as shown in the example below.

The attribute *cert* may be used to indicate the degree of confidence ascribed by the encoder to the proposed emendation on a broad scale: high, medium, or low. The attribute *resp* is used to indicate who is responsible for the proposed emendation. Its value is a pointer, which will typically indicate a <respStmt> or <name> element in the header of the transcribed document, but can point anywhere, for example to some online authority file. Using these two attributes, the <corr> element presented above might usefully be enhanced as follows:

```
<!-- somewhere in the header ... --><name  xml:id="ETD">E Talbot Donaldson</name>
<!-- ... --> And of so parfit wis a <choice>
  <sic>wight</sic>
  <corr  resp="#ETD"  cert="medium">wright</corr>
</choice> ywroght?
```

As remarked above, where the same annotation applies to several corrections, this may be represented by supplying multiple pointers on the note. Consider for example such corrections as the following, in Dudo of S. Quentin. Parkes cites two cases in this manuscript of the same phenomenon:

```
quamuis <choice  xml:id="sic-1">
  <sic>mens</sic>
  <corr>iners</corr>
</choice> que nutu dei gesta
sunt ... unde esset uiriliter
<choice  xml:id="sic-2">
  <corr>uegetata</corr>
  <sic>negata</sic>
</choice>
```

which may be described as follows:

```
<note target="#sic-1 #sic-2">Substitution of a more familiar word which resembles
graphically what the scribe should be copying but which does not make sense in
the context.</note>
```

The *target* attribute on the <note> element indicates the <choice> elements which exemplify this kind of scribal error. This necessitates the addition of an identifier to each <choice> element. However, if the number of corrections is large and the number of notes is small, it may well be both more practical and more appropriate to regard the collection of annotations as constituting a typology and then use the *type* attribute. Suppose that the note given above is one of half a dozen possible kinds of corrected phenomena identified in a given text; others might include, say, 'repetition of a word from the preceding line', etc. The *type* attribute on the <corr> element can be used to specify an arbitrary code for the particular kind of correction (or other editorial intervention) identified within it. This code can be chosen freely and is not treated as a pointer.

```
quamuis <choice>
  <sic>mens</sic>
  <corr type="graphSubs">iners</corr>
</choice> que nutu dei gesta sunt ... unde
esset uiriliter
<choice>
  <corr type="graphSubs">uegetata</corr>
  <sic>negata</sic>
</choice>
```

Note that this encoding might be extended to include a range of possible corrections:

```
quamuis <choice>
  <sic>mens</sic>
  <corr type="graphSubs">iners</corr>
  <corr type="reversal">inres</corr>
</choice> que
nutu dei gesta sunt ...
```

In addition, the conscientious encoder will provide documentation explaining the circumstances in which particular codes are judged appropriate. A suitable location for this might be within the <correction> element of the <encodingDesc> of the header, which might include a <list> such as the following:

```
<correction>
  <p>The following codes are used to categorize corrections identified in this
     transcription: <list type="gloss">
      <label>graphSubs</label>
      <item>Substitution of a more familiar word which resembles graphically
            what the scribe should be copying but which does not make sense in the
            context.</item>
<!-- ... -->
    </list>
  </p>
</correction>
```

A *subtype* attribute may be used in conjunction with the *type* for subclassification purposes: the above examples might thus be represented as <choice type="substitution" subtype="graphicResemblance"> for example.

For a given project, it may well be desirable to limit the possible values for the *type* or *subtype* attributes automatically. This is easily done but requires customization of the TEI system using techniques described in *24.3. Customization*, in particular *24.3.1.3. Modification of Attribute and Attribute Value Lists*, which should be consulted for further information on this topic.

When making a correction in a source which forms part of a textual tradition attested by many witnesses, a textual editor will sometimes use a reading from one witness to correct the reading of the source text. In the general case, such encoding is best achieved with the mechanisms provided by the module for textual criticism described in chapter *13. Critical Apparatus*. However, for simple cases, the *source* attribute of the <corr> element may suffice. In the passage from Chaucer's *Wife of Bath's Tale* mentioned above, Parkes proposes to emend the problematic word *wight* to *wyf* which is the reading found in the Cambridge manuscript Gg.1. 27. This may be simply represented as follows:

```
 And of so
parfit wis a <choice>
  <sic>wight</sic>
  <corr  resp="#mp"  source="#Gg">wyf</corr>
</choice> ywroght?
```

The value of the *source* attribute here is, like the value of the *resp* attribute, a pointer, in this case indicating the manuscript used as a witness. Elsewhere in the transcribed text, a list of witnesses used in this text will be given, one of which has an identifier `Gg`. Each witness will be represented either by a <witness> element (see *13.1. The Apparatus Entry, Readings, and Witnesses*) or more fully by an <msDesc> element (see *11. Manuscript Description*):

```
<msDesc  xml:id="Gg">
  <msIdentifier>
    <settlement>Cambridge</settlement>
    <repository>University Library</repository>
    <idno>Gg.1. 27</idno>
  </msIdentifier>
<!-- further description of the manuscript here -->
</msDesc>
```

The <app> element described in chapter *13. Critical Apparatus* provides a more powerful way of representing all three possible readings in parallel:

```
And of so parfit wis a <app>
  <rdg  wit="#Hg">wight</rdg>
  <rdg  wit="#Ln #Ry2 #Ld">wright</rdg>
  <rdg  wit="#Gg">wyf</rdg>
</app>
```

This encoding simply records the three readings found in the various traditions, and gives (by means of the *wit* attribute) an indication of the witnesses supporting each. If the *resp* attribute were supplied on the <rdg> element, it would indicate the person responsible for asserting that the manuscript indicated has this reading, who is not necessarily the same as the person responsible for asserting that this reading should be used to correct the others. Editorial intervention elements such as <corr> can however be nested within a <rdg> to provide this additional information:

```
And of so parfit wis a <app>
  <rdg  wit="#Hg">wight</rdg>
  <rdg  wit="#Ln #Ry2 #Ld">
```

```
    <corr resp="#ETD">wright</corr>
  </rdg>
  <rdg wit="#Gg">
    <corr resp="#mp">wyf</corr>
  </rdg>
</app>
```

This encoding asserts that the reading *wyf* found in Gg is regarded as a correction by Parkes.

Like the *resp* attribute, the *cert* attribute may be used with both <corr> and <rdg> elements. When used on the <rdg> element, these attributes indicate confidence in and responsibility for identifying the reading within the sources specified; when used on the <corr> element they indicate confidence in and responsibility for the use of the reading to correct the base text. If no other source is indicated (either by the *source* attribute, or by the *wit* attribute of a parent <rdg>), the reading supplied within a <corr> has been provided by the person indicated by the *resp* attribute.

If it is desired to express certainty of or responsibility for some other aspect of the use of these elements, then the mechanisms discussed in chapter *22. Certainty, Precision, and Responsibility* may be found useful. See also *12.3.2.2. Hand, Responsibility, and Certainty Attributes* for further discussion of the issues of certainty and responsibility in the context of transcription.

#### 12.3.1.4 Additions and Deletions

Additions and deletions observed in a source text may be described using the following elements:

`<add>` (addition) contains letters, words, or phrases inserted in the source text by an author, scribe, or a previous annotator or corrector.

`<addSpan>` (added span of text) marks the beginning of a longer sequence of text added by an author, scribe, annotator or corrector (see also <add>).

`<del>` (deletion) contains a letter, word, or passage deleted, marked as deleted, or otherwise indicated as superfluous or spurious in the copy text by an author, scribe, or a previous annotator or corrector.

`<delSpan>` (deleted span of text) marks the beginning of a longer sequence of text deleted, marked as deleted, or otherwise signaled as superfluous or spurious by an author, scribe, annotator, or corrector.

Of these, <add> and <del> are included in the core module, while <addSpan> and <delSpan> are available only when using the module defined in this chapter. These particular elements are members of the `att.spanning` class, from which they inherit the following attribute:

`att.spanning` provides attributes for elements which delimit a span of text by pointing mechanisms rather than by enclosing it.

- `@spanTo` indicates the end of a span initiated by the element bearing this attribute.

Further characteristics of each addition and deletion, such as the hand used, its effect (complete or incomplete, for example), or its position in a sequence of such operations may conveniently be recorded as attributes of these elements, all of which are members of the `att.transcriptional` class:

`att.transcriptional` provides attributes specific to elements encoding authorial or scribal intervention in a text when transcribing manuscript or similar sources.

- `@seq` (sequence) assigns a sequence number related to the order in which the encoded features carrying this attribute are believed to have occurred.
- `@status` indicates the effect of the intervention, for example in the case of a deletion, strikeouts which include too much or too little text, or in the case of an addition, an insertion which duplicates some of the text already present.
- `@hand [att.written]` points to a <handNote> element describing the hand considered responsible for the content of the element concerned.

As described in section *3.5. Simple Editorial Changes*, the <add> element is used to record any manuscript addition observed in the text, whether it is considered to be authorial or scribal. In the autograph manuscript of Max Beerbohm's *The Golden Drugget*, the author's addition of *do ever* may be recorded as follows, with the *hand* attribute indicating that the addition was Beerbohm's by referencing a <handNote> element defined elsewhere in the document (see further *12.3.2.1. Document Hands*):

```
Some things are best
at first sight. Others — and here is one of them — <add  hand="#mb">do ever</add>
improve by recognition [...]
<handNote  xml:id="mb">Max Beerbohm
holograph</handNote>
```

The <del> element is used to record manuscript deletions in a similar way. In the autograph manuscript of D. H. Lawrence's *Eloi, Eloi, lama sabachthani* the author's deletion of *my* may be recorded as follows. In this case, the *hand* attribute indicating that the deletion was Lawrence's is complemented by a *rend* attribute indicating that the deletion was by strike-through:

```
For I hate this <del  rend="strikethrough"  hand="#dhl">my</del> body, which is so dear
to me ...

<handNote  xml:id="dhl">D H Lawrence holograph</handNote>
```

If deletions are classified systematically, the *type* attribute may be useful to indicate the classification; when they are classified by the manner in which they were effected, or by their appearance, however, this will lead to a certain arbitrariness in deciding whether to use the *type* or the *rend* attribute to hold the information. In general, it is recommended that the *rend* attribute be used for description of the appearance or method of deletion, and that the *type* attribute be reserved for higher level or more abstract classifications.

The *place* attribute is also available to indicate the location of an addition. For example, consider the following passage from a draft letter by Robert Graves: At the end of this extract, the writer inserts the word 'cant,' above the

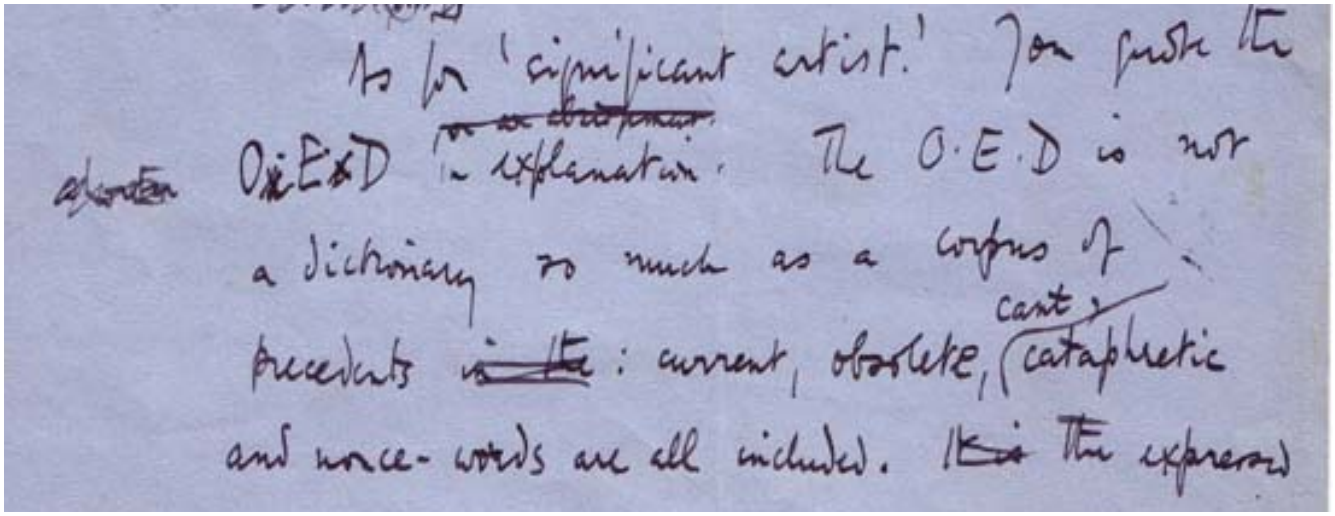



Figure 21: Draft letter from Robert Graves to Desmond Flower, 17 Dec 1938 (detail).

line, with a stroke to indicate insertion. Assuming that we have previously defined the identifier RG somewhere:

```
<listPerson>
  <person  xml:id="RG">
<!-- information about Robert Graves here -->
  </person>
</listPerson>
```

, this extract might now be encoded as follows:

```
 The O.E.D. is not a
dictionary so much as a corpus of precedents <del  hand="#RG">in the</del>:
current, obsolete, <add  hand="#RG"  place="above">cant,</add> cataphretic and
nonce-words are all included.
```

A little earlier in the same extract, Graves writes 'for an abridgement' above the line, and then deletes it. This may be encoded similarly:

```
 As for 'significant artist.' You quote the O.E.D <del>
  <add  hand="#RG"  place="above">for an abridgement</add>
</del> in explanation...
```

Similarly, in the margin, the word 'Norton' has been added and then deleted:

```
 You
quote the <del>
  <add  hand="#RG"  place="margin">Norton</add>
</del> O.E.D...
```

The word 'O.E.D.' in this first sentence has also clearly been the result of some redrafting: it may be that Graves started to write 'Oxford', and then changed it; it may be that he inserted other punctuation marks between the letters before replacing them with the centre dots used elsewhere to represent this acronym. We do not deal with these possibilities here, and mention them only to indicate that any encoding of manuscript material of this complexity will need to make decisions about what is and is not worth mentioning.

An encoder may also wish to indicate that an addition replaces a specific deletion, that is to encode a substitution as a single intervention in the text. This may be achieved by grouping the addition and deletion together within a <subst> element. At the end of the passage illustrated above, Graves first writes 'It is the expressed…', then deletes 'It is', and substitutes an uppercase T at the start of 'the'.

```
 ... are all included. <del  hand="#RG">It is</del>
<subst>
  <add>T</add>
  <del>t</del>
</subst>he expressed
```

The use of this element and of the *seq* attribute to indicate the order in which interventions such as deletions are believed to have occurred are further discussed in section *12.3.1.5. Substitutions* below.

The <add> and <del> elements defined in the core module suffice only for the description of additions and deletions which fit within the structure of the text being transcribed, that is, which each deletion or addition is completely contained by the structural element (paragraph, line, division) within which it occurs. Where this is not the case, for example because an individual addition or deletion involves several distinct structural subdivisions, such as poems or prose items, or otherwise crosses a structural boundary in the text being encoded, special treatment is needed. The <addSpan> and <delSpan> elements are provided by this module for that purpose. (For a general discussion of the issue see further *21. Non-hierarchical Structures*).

In this example of the use of <addSpan>, the insertion by Helgi Ólafsson of a gathering containing four neo-Eddic poems into *Lbs 1562 4to* is recorded as follows. A <handNote> element is first declared, within the header of the document, to associate the identifier heol with Helgi. Each of the added poems is encoded as a distinct  element. In the body of the text, an <addSpan> element is placed to mark the beginning of the span of added text, and an <anchor> is used to mark its end. The *hand* attribute on the <addSpan> element ascribes responsibility for the addition to the manuscript to Helgi, and the *spanTo* attribute points to the end of the added text:

```
<handNote  xml:id="heol"
  scribe="HelgiÓlafsson"/>
<!-- ... -->
<body>
  <div>
<!-- text here -->
  </div>
  <addSpan  n="added gathering"  hand="#heol"
    spanTo="#p025"/>
  <div>
<!-- text of first added poem here -->
  </div>
  <div>
<!-- text of second added poem here -->
  </div>
  <div>
<!-- text of third added poem here -->
  </div>
  <div>
<!-- text of fourth added poem here -->
  </div>
  <anchor  xml:id="p025"/>
  <div>
<!-- more text here -->
  </div>
</body>
```

The <delSpan> element is used in the same way. An authorial manuscript will often contain several occasions where sequences of whole lines are marked for deletion, either by boxes or by being struck out. If the encoder is marking up individual verse lines with the <l> element, such deletions are problematic: deletion of two consecutive lines should be regarded as a single deletion, but the <del> element must be properly nested within a single <l> element. The <delSpan> element solves this problem:

```
<l>Flowed up the hill and down King William Street,</l>
<delSpan  spanTo="#EPdelEnd"  resp="#EP"
  rend="strikethrough"/>
<l>To where Saint Mary Woolnoth kept the time,</l>
<l>With a dead sound on the final stroke of nine.</l>
<anchor  xml:id="EPdelEnd"/>
<l>There I saw one I knew, and stopped him, crying "Stetson!</l>...
```

It is also often the case that deletions and additions may themselves contain other deletions and additions. For example, in Thomas Moore's autograph of the second version of *Lalla Rookh* two lines are marked for omission by vertical strike-through. Within the first of the two lines, the word *upon* has also been struck out, and the word *over* has been added:

```
<l>
  <delSpan  rend="verticalStrike"
    spanTo="#delend01"/> Tis moonlight
<del>upon</del>
  <add>over</add> Oman's sky
</l>
<l>Her isles of pearl look lovelily<anchor  xml:id="delend01"/>
</l>
```

In this case the <anchor> and <delSpan> have been placed within the structural elements (the <l>s) rather than between, as in the previous example. This is to indicate that placement of these empty elements is arbitrary.

The text deleted must be at least partially legible, in order for the encoder to be able to transcribe it. If all of part of it is not legible, the <gap> element should be used to indicate where text has not been transcribed, because it could not be. The <unclear> element described in section *12.3.3.1. Damage, Illegibility, and Supplied Text* may be used to indicate areas of text which cannot be read with confidence. See further section *12.3.1.7. Text Omitted from or Supplied in the Transcription* and section *12.3.3.1. Damage, Illegibility, and Supplied Text.*

#### 12.3.1.5 Substitutions

Substitution of one word or phrase for another is perhaps the most common of all phenomena requiring special treatment in transcription of primary textual sources. It may be simply one word written over the top of another, or deletion of one word and its replacement by another written above it by the same hand on the same occasion; the deletion and replacement may be done by different hands at different times; there may be a long chain of substitutions on the same stretch of text, with uncertainty as to the order of substitution and as to which of many possible readings should be preferred.

As we have shown, the simplest method of recording a substitution is simply to record both the addition and the deletion. However, when the module defined by this chapter is in use, additional elements are available to indicate that the encoder believes the addition and the deletion to be part of the same intervention: a substitution.

`<subst>` (substitution) groups one or more deletions (or surplus text) with one or more additions when the combination is to be regarded as a single intervention in the text.

`<substJoin>` (substitution join) identifies a series of possibly fragmented additions, deletions, or other revisions on a manuscript that combine to make up a single intervention in the text.

Using the <subst> element, the example at the end of the last section might be encoded as follows:

```
<l>
  <delSpan  rend="verticalStrike"
    spanTo="#delend02"/> Tis moonlight
<subst>
    <del>upon</del>
    <add>over</add>
  </subst> Oman's sky
</l>
<l>Her isles of pearl look lovelily<anchor  xml:id="delend02"/>
</l>
```

Since the purpose of this element is solely to group its child elements together, the order in which they are presented is not significant. When both deletion and addition are present, it may not always be clear which occurs first: using the *seq* attribute is a simple way of resolving any such ambiguities.

For example, returning to the example from William James, in a passage first written out by James as 'One must have lived longer with this system, to appreciate its advantages', the word *this* is first replaced by *such a* and this is then replaced by *a*.[80] This may be encoded as follows, representing the two changes as a sequence of additions and deletions:

```
One must have lived longer
with <subst>
  <del  seq="1">this</del>
  <del  seq="2">
    <add  seq="1">such a</add>
  </del>
  <add  seq="2">a</add>
</subst> system, to appreciate its advantages.
```

[80] The manuscript contains several other substitutions, ignored here for the sake of clarity.

Note the nesting of an <add> element within a <del> to record text first added, then deleted in the source. The numbers assigned by the *seq* attribute may be used to identify the order in which the various additions and deletions are believed by the encoder to have been carried out, and thus provide a simple method of supporting the kind of 'genetic' textual criticism typified by (for example) Hans Walter Gabler's work on the reconstruction of the 'overlay' levels implicit in the manuscripts of James Joyce's *Ulysses*. A fuller and more complex way of supporting such an approach is discussed in *12.7. Identifying Changes and Revisions*.

A special case of a substitution may consist of a superfluous word or phrase that is silently replaced by some addition. E.g. a scribe abandons a word (without indicating it should be deleted), and then writes a replacement word immediately after. Here the encoder may interpret this as an 'unmarked' deletion which can then be combined with a corresponding addition to a substitution.

The case of a single substitution or scribal correction that involves non-contiguous addition and deletion can be handled by using the <substJoin> element to make an explicit connection between one or more <add> and <del> elements. In the following example from Thomas Moore's Lalla Rookh, the deletion and addition are not contiguous: they are separated by the word 'thus', which is not part of the scribal intervention being marked. Because of this intervening text, it would be inappropriate to use <subst> to group this <add> and <del>. <substJoin> allows the encoder to indicate that additions and deletions separated in this way are part of a single scribal intervention:

```
 While <del  xml:id="change1">pondering</del> thus <add  xml:id="change2">she
mus'd</add>, her pinions
fann'd
<substJoin  target="#change1 #change2"/>
```

Note that, unlike <subst>, the placement of the <substJoin> is arbitrary. It may occur before or after the relevant <add> and <del> elements.

As a more complex example, consider the following passage: This passage might be encoded as follows:

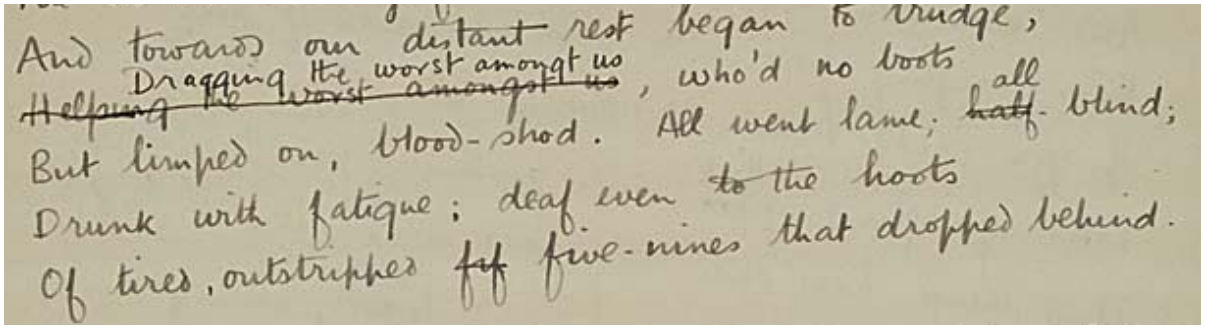



Figure 22: Detail from *Dulce et decorum est* autograph manuscript in the English Faculty Library, Oxford University.

```
<l>And towards our distant rest began to trudge,</l>
<l>
  <subst>
    <del>Helping the worst amongst us</del>
    <add>Dragging the worst amongt
        us</add>
  </subst>, who'd no boots
</l>
<l>But limped on, blood-shod. All went lame; <subst>
    <del  status="shortEnd">half-</del>
    <add>all</add>
  </subst> blind;</l>
<l>Drunk with fatigue ; deaf even to the hoots</l>
```

```
<l>Of tired, outstripped <del>fif</del> five-nines that dropped
behind.</l>
```

In this representation,

- the authorial slip (*amongt* for *amongst*) is retained without comment.
- the other two authorial corrections are marked as substitutions, each combining a deletion and an addition.
- the false start *fif* in the last line is simply marked as a deletion;

The <app> element presented in chapter *13. Critical Apparatus* provides similar facilities, by treating each state of the text as a distinct reading. The <rdg> element has a *varSeq* attribute which may be used in the same way as the *seq* attribute to indicate the preferred sequence. The James example above might thus be represented as follows:

```
One must have lived longer with <app>
  <rdg  varSeq="1">
    <del>this</del>
  </rdg>
  <rdg  varSeq="2">
    <del>
      <add>such a</add>
    </del>
  </rdg>
  <rdg  varSeq="3">
    <add>a</add>
  </rdg>
</app> system, to appreciate its advantages.
```

#### 12.3.1.6 Cancellation of Deletions and Other Markings

An author or scribe may mark a word or phrase in some way, and then on reflection decide to cancel the marking. For example, text may be marked for deletion and the deletion then cancelled, thus restoring the deleted text. Such cancellation may be indicated by the <restore> element:

`<restore>` (restore) indicates restoration of text to an earlier state by cancellation of an editorial or authorial marking or instruction.

This element bears the same attributes as the other transcriptional elements. These may be used to supply further information such as the hand in which the restoration is carried out, the type of restoration, and the person responsible for identifying the restoration as such, in the same way as elsewhere.

Presume that Lawrence decided to restore *my* to the phrase of *Eloi, Eloi, lama sabachthani* first written 'For I hate this my body', with the *my* first deleted then restored by writing 'stet' in the margin. This may be encoded:

```
For I hate this <restore  hand="#dhl"
  type="marginalStetNote">
  <del>my</del>
</restore> body
```

Another feature commonly encountered in manuscripts is the use of circles, lines, or arrows to indicate transposition of material from one point in the text to another. No specific markup for this phenomenon is proposed at this time. Such cases are most simply encoded as additions at the point of insertion and deletions at the point of encirclement or other marking.

#### 12.3.1.7 Text Omitted from or Supplied in the Transcription

Where text is not transcribed, whether because of damage to the original, or because it is illegible, or for some other reason such as editorial policy, the <gap> core element may be used to register the omission; where such text is transcribed, but the editor wishes to indicate that they consider it to be superfluous, for example because it

is an inadvertent scribal repetition or an interpolation from another source, the <surplus> element may be used in preference. Where the editor believes text to be interpolated but genuine, the <secl> element may be used instead. Where text not present in the source is supplied (whether conjecturally or from other witnesses) to fill an apparent gap in the text, the <supplied> element may be used.

`<gap>` (gap) indicates a point where material has been omitted in a transcription, whether for editorial reasons described in the TEI header, as part of sampling practice, or because the material is illegible, invisible, or inaudible.

`@reason` (reason) gives the reason for omission.

`@agent` (agent) in the case of text omitted because of damage, categorizes the cause of the damage, if it can be identified.

`<surplus>` (surplus) marks text present in the source which the editor believes to be superfluous or redundant.

`@reason` one or more words indicating why this text is believed to be superfluous, e.g. *repeated*, *interpolated* etc.

`<secl>` (secluded text) Secluded. Marks text present in the source which the editor believes to be genuine but out of its original place (which is unknown).

`@reason` one or more words indicating why this text has been secluded, e.g. *interpolated* etc.

`<supplied>` (supplied) signifies text supplied by the transcriber or editor for any reason; for example because the original cannot be read due to physical damage, or because of an obvious omission by the author or scribe.

`@reason` one or more words indicating why the text has had to be supplied, e.g. *overbinding*, *faded-ink*, *lost-folio*, *omitted-in-original*.

By its nature, the <gap> element has no content. It marks a point in the text where nothing at all can be read, whether because of authorial or scribal erasure, physical damage, or any other form of illegibility. Its attributes allow the encoder to specify the amount of text which is illegible in this way at this point, using any convenient units, where this can be determined. For example, in the Beerbohm manuscript of *The Golden Drugget* cited above, the author has erased a passage amounting about 10 cm in length by inking over it completely:

```
Others <gap  reason="cancelled"  quantity="10"
  unit="cm"/>—and here is one of
them...
```

In an autograph letter of Sydney Smith now in the Pierpont Morgan library three words in the signature are quite illegible:

```
I am dr Sr yr <gap  reason="illegible"  quantity="3"
  unit="word"/>Sydney Smith
```

The degree of precision attempted when measuring the size of a gap will vary with the purpose of the encoding and the nature of the material: no particular recommendation is made here.

As noted above, the <gap> element should only be used where text has not been transcribed. If partially legible text has been transcribed, one of the elements <damage> and <unclear> should be used instead (these elements are described in section *12.3.3.1. Damage, Illegibility, and Supplied Text*); if the text is legible and has been transcribed, but the editor wishes to indicate that they regard it is superfluous or redundant, then the element <surplus> may be used in preference to the core element <sic> used to indicate text regarded as erroneous.

Amongst the many examples cited in Hans Krummrey & Silvio Panciera's classic text on the editing of epigraphic inscriptions is the following. In a late classical inscription, the form 'dedikararunt' is encountered. The editor may choose any of the following three possibilities:

- mark this as an erroneous form

```
<sic>dedikararunt</sic>
```

- additionally supply a corrected form

```
<choice>
  <sic>dedikararunt</sic>
  <corr>dedikarunt</corr>
</choice>
```

- indicate that the erroneous form contains surplus characters which the editor wishes to suppress

```
dedika<surplus>ra</surplus>runt
```

The <surplus> element may also be used to mark up interpolations, as in the following example taken from a 13th century Italian source:

```
<l  n="4">a darmi morte, poi m'avete preso <surplus  reason="interpolated">a
    tradimento</surplus>
</l>
<l  n="5">sì com' l'uccellator prende l'uccello</l>
<gap/>
<l  n="43">e lettere dintorno che diriano <surplus  reason="interpolated">in questa
    guisa</surplus>
</l>
<l  n="44">Più v'amo, dëa, che non faccio Deo</l>
```

The words marked as <surplus> here are metrically inconsistent with the rest and have been marked by the editor as such.

In the case of an interpolation which the editor regards as genuine (i.e. written by the author in question), but out of its original place, the <secl> element should be used instead of <surplus>. For example:

```
<sp>
  <ab>
    <lb  n="545"/>Great praise and thanks be to Perfidy as she
    <lb  n="546"/>deserves, since by our swindles, tricks, and clever moves, relying
    <lb  n="547"/>on the daring of our shoulder blades and the excellence of our
    <lb  n="548"/>forearms who went against cattle-prods, hot iron-blades,
    <lb  n="549-550"/>crosses and shackles, neck-irons, chains, prisons, collars, fetters,
    <lb  n="551"/>and yokes, the fiercest painters fully acquainted with our backs
    <lb  n="552"/><secl>who have often before put scars on our shoulder blades</secl>
    ...
  </ab>
</sp>
```

The final line is bracketed in the Loeb edition, with a note: ‘versum secl. Bothe’, meaning Bothe regarded this line as Plautine, but probably interpolated. It is easy to see how the line might have crept in as a gloss on the metaphor in the previous line.

If some part of the source text is completely illegible or missing, an encoder may sometimes wish to supply new (conjectural) material to replace it. This conjectural reading is analogous to a correction in that it contains text provided by the encoder and not attested in the source. This is not however a correction, since no error is necessarily present in the original; for that reason a different element <supplied> should be used. If another (imaginary) copy

of the letter above preserved the signature as reading 'I am dear Sir your very humble Servt Sydney Smith', the text illegible in the autograph might be supplied in the transcription:

```
I am dr Sr yr
<supplied  reason="illegible"  resp="#msm"
  source="#Ry2">very humble
Servt</supplied> Sydney Smith
```

Here the *source* and *resp* attributes are used, as elsewhere, to indicate respectively the sigil of a manuscript from which the supplied reading has been taken, and the identifier of the person responsible for deciding to supply the text. If the *source* attribute is not supplied, the implication is that the encoder (or whoever is indicated by the value of the *resp* attribute) has supplied the missing reading. Both <gap> and <supplied> may be used in combination with <unclear>, <damage>, and other elements; for discussion, see section *12.3.3.2. Use of the gap, del, damage, unclear, and supplied Elements in Combination.*

## 12.3.2 Hands and Responsibility

This section discusses in more detail the representation of aspects of responsibility perceived or to be recorded for the writing of a primary source. These include points at which one scribe takes over from another, or at which ink, pen, or other characteristics of the writing change. A discussion of the usage of the *hand*, *resp*, and *cert* attributes is also included.

### 12.3.2.1 Document Hands

For many text-critical purposes it is important to signal the person responsible (the *hand*) for the writing of a whole document, a stretch of text within a document, or a particular feature within the document. A hand, as the name suggests, need not necessarily be identified with a particular known (or unknown) scribe or author; it may simply indicate a particular combination of writing features recognized within one or more documents. The examples given above of the use of the *hand* attribute with coding of additions and deletions illustrate this.

The <handNote> element is used to provide information about each hand distinguished within the encoded document.

`<handNote>` (note on hand) describes a particular style or hand distinguished within a manuscript.

A <handNote> element, with an identifier given by its *xml:id* attribute, may appear in either of two places in the TEI header, depending on which modules are included in a schema. When the `transcr` module defined by the present chapter is used, the element <handNotes> is available, within the <profileDesc> element of the TEI header, to hold one or more <handNote> elements. When the `msdescription` module defined in chapter *11. Manuscript Description* is included, the <handDesc> element described in *11.7.2. Writing, Decoration, and Other Notations* also becomes available as part of a structured manuscript description. The encoder may choose to place <handNote> elements identifying individual hands in either location without affecting their accessibility since the element is always addressed by means of its *xml:id* attribute. The <handDesc> element may be more appropriate when a full cataloguing of each manuscript is required; the <handNotes> element if only a brief characterization of each hand is needed. It is also possible to use the two elements together if, for example, the <handDesc> element contains a single summary describing all the hands discursively, while the <handNotes> element gives specific details of each. The choice will depend on individual encoders' priorities.

As shown above, the *hand* attribute is available on several elements to indicate the hand in which the content of the element (usually a deletion or addition) is carried out. The <handShift> element may also be used within the body of a transcription to indicate where a change of hand is detected for whatever reason.

`<handShift>` (handwriting shift) marks the beginning of a sequence of text written in a new hand, or the beginning of a scribal stint.

A <handShift> element can be used to indicate a change of hand even within an element with a *hand* attribute. The text following the <handShift> must be considered to be in the new hand.

Both <handShift> and <handNote> are members of the `att.handFeatures` class, and thus share the following attributes:

`att.handFeatures` provides attributes describing aspects of the hand in which a manuscript is written.

`@scribe` gives a name or other identifier for the scribe believed to be responsible for this hand.

`@script` characterizes the particular script or writing style used by this hand, for example *secretary*, *copperplate*, *Chancery*, *Italian*, etc.

`@scribeRef` points to a full description of the scribe concerned, typically supplied by a <person> element elsewhere in the description.

`@scriptRef` points to a full description of the script or writing style used by this hand, typically supplied by a <scriptNote> element elsewhere in the description.

`@medium` describes the tint or type of ink, e.g. *brown*, or other writing medium, e.g. *pencil*.

`@scope [att.scope]` indicates the scope of application of the element

A single hand may employ different writing styles and inks within a document, or may change character. For example, the writing style might shift from 'anglicana' to 'secretary', or the ink from blue to brown, or the character of the hand may change. Simple changes of this kind may be indicated by assigning a new value to the appropriate attribute within the <handShift> element. It is for the encoder to decide whether a change in these properties of the writing style is so marked as to require treatment as a distinct hand.

Where such a change is to be identified, the *new* attribute indicates the hand applicable to the material following the <handShift>. The sequence of such <handShift> elements will often, but not necessarily, correspond with the order in which the material was originally written. Where this is not the case, the facilities described in section *12.7. Identifying Changes and Revisions* may be found helpful.

As might be expected, a single hand may also vary renditions within the same writing style, for example medieval scribes often indicate a structural division by emboldening all the words within a line. Such changes should be indicated by use of the *rend* attribute, in the same manner as underlining, emboldening, font shifts, etc. are represented in transcription of a printed text, rather than by introducing a new <handShift> element.

In the following example there is a change of ink within a single hand. This is simply indicated by a new value for the *medium* attribute on the <handShift> element:

```
<l>When wolde the cat dwelle in his ynne</l>
<handShift  medium="greenish-ink"/>
<l>And if the cattes skynne be slyk <handShift  medium="black-ink"/> and
gaye</l>
```

In the following example, the encoder has identified two distinct hands within the document and given them identifiers h1 and h2, by means of the following declarations included in the document's TEI header:

```
<handNotes>
  <handNote  xml:id="h1"  script="copperplate"
    medium="brown-ink">Carefully written
     with regular descenders</handNote>
  <handNote  xml:id="h2"  script="print"
    medium="pencil">Unschooled
     scrawl</handNote>
</handNotes>
```

Then the change of hand is indicated in the text:

```
<handShift  new="#h1"  resp="#das"/>... and that good Order Decency and regular worship
may be once
more introduced and Established in this Parish according to the Rules and
Ceremonies of the Church of England and as under a good Consciencious and sober
Curate there would and ought to be <handShift  new="#h2"  resp="#das"/> and for that
purpose the parishioners pray
```

When a more precise or nuanced discussion of the writing in a manuscript is required, the <handNote> and <scriptNote> elements discussed in *11.7.2. Writing, Decoration, and Other Notations* should be used. Either element may serve as the target for a <handShift>.

#### 12.3.2.2 Hand, Responsibility, and Certainty Attributes

The *hand* and *resp* attributes have similar, but not identical, meanings. Observe their distinctive uses in the following encoding of the William James passage mentioned above in section *12.3.1.3. Correction and Conjecture*. In this example, the *But* inserted by James is tagged as an <add>, and the consequent editorial correction of *One* to *one* treated separately:

```
<add place="above" resp="#FB" hand="#WJ">But</add>
<choice>
  <sic>One</sic>
  <corr resp="#FB">one</corr>
</choice> must have lived ...
<!-- elsewhere -->
<respStmt xml:id="FB">
  <resp>editorial changes</resp>
  <name>Fredson Bowers</name>
</respStmt>
<respStmt xml:id="WJ">
  <resp>authorial changes</resp>
  <name>William James</name>
</respStmt>
```

As in this example, *hand* should be reserved for indicating the hand of any form of marking—here, addition but also deletion, correction, annotation, underlining, etc.—within the primary text being transcribed. The scribal or authorial responsibility for this marking may be inferred from the value of the *hand* attribute. The value of the *hand* attribute should be a pointer to a hand identifiers typically declared in the document header but potentially in another document or repository (see section *12.3.2.1. Document Hands*).

The *resp* attribute, by contrast, indicates the person responsible for deciding to mark up this part of the text with this particular element. In the case of the <add> element, for example, the *resp* attribute will indicate the responsibility for identifying that the addition is indeed an addition, and also (if the *hand* attribute is supplied) to which hand it should be attributed. In this case, Bowers is credited with identifying the hand as that of William James. In the case of the <corr> element, the *resp* attribute indicates who is responsible for supplying the intellectual content of the correction reported in the transcription: here, Bowers' correction of 'One' to 'one'. In the case of a deletion, the *resp* attribute will similarly indicate who bears responsibility for identifying or categorizing the deletion itself, while other attributes (*hand* most obviously) attribute responsibility for the deletion itself. It should be noted that the *source* attribute may be used in a simiilar fashion to indicate, for example, when an encoding decision is based on the work of a previous editor or on an article. In that case, the *source* would point to a <bibl> in the bibliography.

In cases where both the *resp* and *cert* attributes are defined for a particular element, the two attributes refer to the same aspect of the markup. The one indicates who is intellectually responsible for some item of information, the other indicates the degree of confidence in the information. Thus, for a correction, the *resp* attribute signifies the person responsible for supplying the correction, while the *cert* attribute signifies the degree of editorial confidence felt in that correction. For the expansion of an abbreviation, the *resp* attribute signifies the person responsible for supplying the expansion and the *cert* attribute signifies the degree of editorial confidence felt in the expansion.

This close definition of the use of the *resp* and *cert* attributes with each element is intended to provide for the most frequent circumstances in which encoders might wish to make unambiguous statements regarding the responsibility for and certainty of aspects of their encoding. The *resp* and *cert* attributes, as so defined, give a convenient mechanism for this. However, there will be cases where it is desirable to state responsibility for and certainty concerning other aspects of the encoding. For example, one may wish in the case of an apparent addition to state the responsibility for the use of the <add> element, rather than the responsibility for identifying the hand of the addition. It may also be that one editor may make an electronic transcription of another editor's printed transcription of a manuscript text—here, one will wish to assign layers of responsibility, so as to allow the reader to determine exactly what in the final transcription was the responsibility of each editor. In these complex cases of divided editorial responsibility for and certainty concerning the content, attributes, and application of a particular element, the more general mechanisms for representing certainty and responsibility described in chapter *22. Certainty, Precision, and Responsibility* should be used.

It should be noted that the certainty and responsibility mechanisms described in chapter *22. Certainty, Precision, and Responsibility* replicate all the functions of the *resp* and *cert* attributes on particular elements. For example, the encoding of Donaldson's conjectured emendation of *wight* to *wright* in line 117 of Chaucer's *Wife of Bath's Prologue* (see *12.3.1.3. Correction and Conjecture*) may be encoded as follows using the *resp* and *cert* attributes on the <corr> element:

```
<choice>
  <sic>wight</sic>
  <corr  resp="#ETD"  cert="medium">wright</corr>
</choice>
```

Exactly the same information could be conveyed using the certainty and responsibility mechanisms, as follows:

```
<choice>
  <corr  xml:id="c117">wright</corr>
  <sic>wight</sic>
</choice>
<certainty  target="#c117"  locus="value"
  degree="0.7"/>
<respons  target="#c117"  locus="value"
  resp="#ETD"/>
```

The choice of which mechanism to use is left to the encoder. In transcriptions where only such statements of responsibility and certainty are made as can be accommodated within the *resp* and *cert* attributes of particular elements, it will be economical to use the *resp* and *cert* attributes of those elements. Where many statements of responsibility and certainty are made which cannot be so accommodated, it may be economical to use the <respons> and <certainty> elements throughout.

The above discussion supposes that in each case an encoder is able to specify exactly what it is that one wishes to state responsibility for and certainty about. Situations may arise when an encoder wishes to make a statement concerning certainty or responsibility but is unable or unwilling to specify so precisely the domain of the certainty or responsibility. In these cases, the <note> element may be used with the *type* attribute set to 'cert' or 'resp' and the content of the note giving a prose description of the state of affairs.

## 12.3.3 Damage and Conjecture

The carrier medium of a primary source may often sustain physical damage which makes parts of it hard or impossible to read. In this section we discuss elements which may be used to represent such situations and give recommendations about how these should be used in conjunction with the other related elements introduced previously in this chapter.

### 12.3.3.1 Damage, Illegibility, and Supplied Text

The <gap> and <supplied> elements described above (section *12.3.1.7. Text Omitted from or Supplied in the Transcription*) should be used with appropriate attributes where the degree of damage or illegibility in a text is such that nothing can be read and the text must be either omitted or supplied conjecturally or from one or more other sources. In many cases, however, despite damage or illegibility, the text may yet be read with reasonable confidence. In these cases, the following elements should be used:

`<damage>` (damage) contains an area of damage to the text witness.

`<damageSpan>` (damaged span of text) marks the beginning of a longer sequence of text which is damaged in some way but still legible.

As members of the class `att.damaged`, these elements bear the following attributes:

`att.damaged` provides attributes describing the nature of any physical damage affecting a reading.

- `@hand [att.written]` points to a <handNote> element describing the hand considered responsible for the content of the element concerned.
- `@agent` categorizes the cause of the damage, if it can be identified.

@*degree* provides a coded representation of the degree of damage, either as a number between 0 (undamaged) and 1 (very extensively damaged), or as one of the codes high, medium, low, or unknown. The <damage> element with the *degree* attribute should only be used where the text may be read with some confidence; text supplied from other sources should be tagged as <supplied>.

@*group* assigns an arbitrary number to each stretch of damage regarded as forming part of the same physical phenomenon.

The class `att.damaged` is a subclass of the class `att.dimensions`, itself a subclass of the class `att.ranging`. Consequently these elements also therefore bear at least the following attributes:

`att.dimensions` provides attributes for describing the size of physical objects.

@*extent* indicates the size of the object concerned using a project-specific vocabulary combining quantity and units in a single string of words.

@*unit* names the unit used for the measurement

@*quantity* specifies the length in the units specified

`att.ranging` provides attributes for describing numerical ranges.

@*min* where the measurement summarizes more than one observation or a range, supplies the minimum value observed.

@*max* where the measurement summarizes more than one observation or a range, supplies the maximum value observed.

@*atLeast* gives a minimum estimated value for the approximate measurement.

@*atMost* gives a maximum estimated value for the approximate measurement.

From the `att.spanning` class, <damageSpan> inherits the following additional attribute:

`att.spanning` provides attributes for elements which delimit a span of text by pointing mechanisms rather than by enclosing it.

@*spanTo* indicates the end of a span initiated by the element bearing this attribute.

The following examples all refer to the recto of folio 5 of the unique manuscript of the Elder Edda. Here, the manuscript of *Vóluspá* has been damaged through irregular rubbing so that letters in various places are obscured and in some cases cannot be read at all.

In the first line of this leaf, the transcriber may believe that the last three letters of *daga* can be read clearly despite the damage:

```
um aldr
d<damage>aga</damage> yndisniota
```

If, as is often the case, the damage crosses structural divisions, so that the <damage> element cannot be nested properly within the containing  elements, the <damageSpan> element may be used, in the same way as the <delSpan> and <addSpan> elements discussed in section *12.3.1.4. Additions and Deletions*.

```
<p>
<!-- ... -->
  <pb  n="5r"/>
  <damageSpan  agent="rubbing"
    extent="whole leaf"  spanTo="#damageEnd"/>
</p>
<p> [...] </p>
<p> [...] <pb  n="5v"  xml:id="damageEnd"/>
</p>
```

Note that in this example the *spanTo* element points to the next <pb> element rather than to an inserted <anchor> element, since it is the whole of the leaf (the text between the two <pb> elements) which has sustained damage. For other techniques of handling non-nesting information, see chapter *21. Non-hierarchical Structures*.

If, as is also likely, the damage affects several disjoint parts of the text, each such part must be marked with a separate <damage> or <damageSpan> element. To indicate that each of these is to be regarded as forming part of the same damaged area, the *group* attribute may be used as in the following example. In this (imaginary) text of Fitzgerald's translation from Omar Khayam, water damage has affected an area covering parts of several lines:

```
<l> The Moving Finger wri<damage  agent="water"  group="1">es; and</damage> having
writ,</l>
<l>Moves <damage  agent="water"  group="1">on: nor all your</damage> Piety nor
Wit</l>
<l>
  <damageSpan  agent="water"  group="1"
    spanTo="#washOut"/>Shall lure it back to
cancel half a Line,
</l>
<l>Nor all your Tears wash <anchor  xml:id="washOut"/> out a Word of it</l>
```

A more general solution to this problem is provided by the <join> element discussed in *17.7. Aggregation* which may be used to link together arbitrary elements of any kind in the transcription. Here, several phenomena of illegibility and conjecture all result from a single cause: an area of damage to the text caused by rubbing at various points. The damage is not continuous, and affects the text at irregular points. In cases such as this, the join element may be used to indicate which tagged features are part of the same physical phenomenon.

If the damage has been so severe as to render parts of the text only imperfectly legible, the <unclear> element should be used to mark the fact. Returning to the Eddic example above, an encoder less confident in the *daga* reading might indicate this as follows:

```
um aldr d<unclear  reason="damage">aga</unclear> yndisniota
```

If it is desired to supply more information about the kind of damage, it is also possible to nest an <unclear> element within the <damage> element:

```
um aldr d<damage  agent="rubbing">
  <unclear>aga</unclear>
</damage> yndisniota
```

Alternatively, the transcriber may not feel able to read the last three letters of *daga* but may wish to supply them by conjecture. Note the use of the *resp* attribute to assign the conjecture to Finnur Jónsson:

```
um aldr d<supplied  reason="rubbing"  resp="#finjon">aga</supplied>
yndisniota
```

The <supplied> element may if desired be enclosed within a <damage> element:

```
um aldr d<damage  agent="rubbing">
  <supplied  source="#msm">aga</supplied>
</damage>
yndisniota
```

The transcriber may also provide alternative conjectures by enclosing multiple <supplied> elements within a <choice> element.

Contrast the use of <gap> in the next line, where the transcriber believes that four letters cannot be read at all because of the damage:

```
þar komr inn dimmi
dreki fliugandi naþr frann neþan <gap  reason="illegible"  agent="rubbing"
  quantity="4"  unit="letter"/>
```

As with <supplied>, this <gap> might be enclosed by a <damage> element.

Where elements are nested in this way, information about agency, etc. is by default inherited. In the following imaginary example, there is a smoke-damaged part within which two stretches can be read with some difficulty, and a third stretch which cannot be read at all:

```
<damage  agent="smoke">
  <unclear>and the proof of this is</unclear>
  <gap/>
  <unclear>margin</unclear>
</damage>
```

The above examples record imperfect legibility due to damage. When imperfect legibility is due to some other reason (typically because the handwriting is ill-formed), the <unclear> element should be used without any enclosing <damage> element. In Robert Southey's autograph of *The Life of Cowper* the final six letters of *attention* are difficult to read because of the haste of the writing, though reasonably certain from the context.

```
and
from time to time invited in like manner his att<unclear>ention</unclear>
```

The *cert* attribute on the <unclear> element may be used to indicate the level of editorial confidence in the reading contained within it.

#### 12.3.3.2 Use of the <gap>, <del>, <damage>, <unclear>, and <supplied> Elements in Combination

The <gap>, <damage>, <unclear>, <supplied>, and <del> elements may be closely allied in their use. For example, an area of damage in a primary source might be encoded with any one of the first four of these elements, depending on how far the damage has affected the readability of the text. Further, certain of the elements may nest within one another. The examples given in the last sections illustrate something of how these elements are to be distinguished in use. This may be formulated as follows:

- where the text has been rendered completely illegible by deletion or damage and no text is supplied by the editor in place of what is lost: place an empty <gap> element at the point of deletion or damage. Note that the <gap> could be wrapped in a <del> or <damage> element. Use the *reason* attribute to state the cause (damage, deletion, etc.) of the loss of text.
- where the text has been rendered completely illegible by deletion or damage and text is supplied by the editor in place of what is lost: surround the text supplied at the point of deletion or damage with the <supplied> element. Use the *reason* attribute to state the cause (damage, deletion, etc.) of the loss of text leading to the need to supply the text.
- where the text has been rendered partly illegible by deletion or damage so that the text can be read but without perfect confidence: transcribe the text and surround it with the <unclear> element. Use the *reason* attribute to state the cause (damage, deletion, etc.) of the uncertainty in transcription and the *cert* attribute to indicate the confidence in the transcription.
- where there is deletion or damage but at least some of the text can be read with perfect confidence: transcribe the text and surround it with the <del> element (for deletion) or the <damage> element (for damage). Use appropriate attribute values to indicate the cause and type of deletion or damage. Observe that the *degree* attribute on the <damage> element permits the encoding to show that a letter, word, or phrase is not perfectly preserved, though it may be read with confidence.

- where there is an area of deletion or damage and parts of the text within that area can be read with perfect confidence, other parts with less confidence, other parts not at all: in transcription, surround the whole area with the <del> element (for deletion; or the <delSpan> element where it crosses a structural boundary); or the <damage> element (for damage). Text within the damaged area which can be read with perfect confidence needs no further tagging. Text within the damaged area which cannot be read with perfect confidence may be surrounded with the <unclear> element. Places within the damaged area where the text has been rendered completely illegible and no text is supplied by the editor may be marked with the <gap> element. For each element, one may use appropriate attribute values to indicate the cause and type of deletion or damage and the certainty of the reading.

The rules for combinations of the <add> and <del> elements, and for the interpretation of such combinations, are similar:

- if one <add> element (with identifier ADD1) contains another (with identifier ADD2), then the addition ADD1 was first made to the text, and later a second addition (ADD2) was made within that added text:

```
This is the text <add  xml:id="ADD1">with some added <add  xml:id="ADD2">(interlinear!)</add>
material</add> as
written.
```

- if one <del> element contains another, and the *seq* attribute does not indicate otherwise, it should be assumed that the inner deletion was made before the enclosing one. In the following example, the word *redundant* was deleted before a second deletion removed the entire passage:

```
<del>This
sentence contains some <del>redundant</del> unnecessary
verbiage.</del>
```

- if a <del> element contains an <add> element, the normal interpretation will be that an addition was made within a passage which was later deleted in its entirety:

```
<del>This
sentence was deleted <add>originally</add> from the
text.</del>
```

- if an <add> element contains a <del> element, the normal interpretation will be that a deletion was made from a passage which had earlier been added:

```
<add>This sentence was
added <del>eventually</del> to the text.</add>
```

- When some text has been blackened out so thoroughly that can no longer be read, the encoding should be:

```
runs out the door in <gap  extent="1 word"/> shirt
```

- For consistency, one might prefer to encode the deletion as such, using <del>, and containing a <gap>, as in the following example:

```
runs out the door in <del>
  <gap  extent="1 word"/>
</del> shirt
```

- This is something that would be necessary if one wanted to encode a <subst> including an illegible deletion:

```
sum<subst>
  <del>
    <gap  extent="2 characters"/>
  </del>
  <add  place="inline">mo</add>
</subst>ns
```

- If some parts of the deleted text are readable, and other parts unreadable, it should be encoded as in the following example:

```
Billy in The <del>
  <gap  extent="1 character"/>ng<gap  extent="2 characters"/>
</del> Silver Dollar..
```

### 12.3.4 Marking up the Writing Process

Modifications of various kinds (correction, addition, deletion, etc.) are frequently found within a single document, and may also be inferred when different documents are compared, although it may be an open question as to whether inter-document discrepancies should be regarded in the same way as intra-document alterations. When two witnesses are collated, we may observe that a word present in one is missing from the other: this does not necessarily imply that the word was added to the first witness, nor that it was deleted from the other.

In this section we discuss a number of elements which may be useful when attempting to record traces of the writing process within a document.

#### 12.3.4.1 Generic Modification

Most, if not all, transcriptional elements imply a certain level of semantic interpretation. For instance, using the <add> element to encode a word or phrase that occupies interlinear space involves a decision that it has been deliberately inserted as an addition rather than an alternative, and indeed a judgment that it was written after, rather than before, the other lines. Where it is felt desirable to keep the recording of 'what is on the page' entirely separate from 'what is the editor's interpretation', the generic <mod> element may be preferred.

`<mod>` represents any kind of modification identified within a single document.

This element simply indicates any kind of modification that has been identified in the document, without prejudice as to its function. Occurrences of the <mod> element may be categorized by means of their *type* attribute, and visual aspects of their appearance can be described by means of the *rend* attribute, but they provide no further interpretation of the function or intention of the passage so marked up. The *spanTo* attribute may be used to indicate the end of a modified passage if this extends across the boundaries of some other XML element, for example from the middle of one line tagged as a <line> to the middle of another <line> some distance further on in the document.

```
<line>words words words <mod  rend="wavy-underlining"
   spanTo="#enduw"/>words with wavy
underline</line> <!-- more lines here -->
<line>wavy underlining finishes
here<anchor  xml:id="enduw"/> more words</line>
```

The distinction between an example such as that above and the simple use of <hi> to mark the visual salience of the underlining (apart from the use of the *spanTo* attribute) is that <hi> does not imply that the visual effect being recorded is understood to represent some kind of modification.

#### 12.3.4.2 Metamarks

By *metamark* we mean marks such as numbers, arrows, crosses, or other symbols introduced by the writer into a document expressly for the purpose of indicating how the text is to be read. Such marks thus constitute a kind of markup of the document, rather than forming part of the text.

`<metamark>` contains or describes any kind of graphic or written signal within a document the function of which is to determine how it should be read rather than forming part of the actual content of the document.

`@function` describes the function (for example status, insertion, deletion, transposition) of the metamark.

`@target` identifies one or more elements to which the metamark applies.

Unlike marginal notes or other additions to the text, metamarks are used by the writer to indicate a deliberate alteration of the writing itself, such as 'move this passage over there'. An addition or annotation by contrast would typically concern some property of the passage other than its intended location or status within the text flow. A metamark may contain text, or some other graphic which the encoder wishes to represent, or it may simply consist of arrows, dots, lines etc. which the encoder simply describes.

The <metamark> element carries a *function* attribute which specifies the function of the metamark, using values such as reorder, flag, delete, insert or used. The passage to which the metamark applies may be indicated in either of two ways: the *target* attribute may be used to point to the element or elements containing the passage concerned, or the *spanTo* element may be used to point to a position in the document at which the passage concerned finishes. In the latter case, the <metamark> itself must be supplied at the position in the document where the passage concerned begins; in the former case it may be supplied at any convenient point. Both attributes should not be supplied.

The following example is taken from an 15th century legal book from the city of Göttingen, containing regulations of everyday life issued by the city council

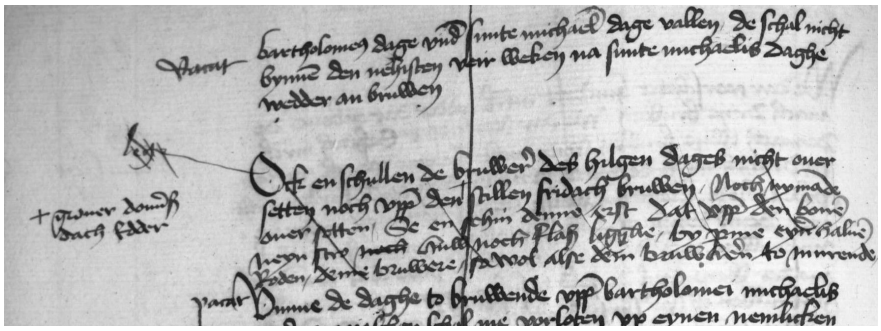


Figure 23: *Kundige bok 2*, fol 1v.

In the second paragraph, the word *lege* ("read") was written in the left hand margin, next to the sentence beginning 'Ock en schullen de bruwere…'. It is thought to function as a metamark, indicating that this sentence forms part of the regulations. A further sentence was then added, while at some later stage the text and also the metamark were deleted. We might encode this as follows as an embedded transcription (keeping in mind that the elements described here can also be used within <text>):

```
<delSpan spanTo="#endDel" change="#L3"/>
<metamark function="flag" target="#zone-1"
  change="#L2">lege</metamark>
<zone xml:id="zone-1" change="#L1">Ock en schullen de bruwere des hilgen dages
nicht over setten noch uppe den stillen fridach bruwen.</zone>
<addSpan spanTo="#endDel" change="#L2"/>
<zone>Noch nymande over setten, se en sehin denne erst, dat uppe den bonen neyn
stro noch, huw noch flaß ligghe, by pine eyner halven roden, deme bruwere so
wol alse dem bruwheren to murende.</zone>
<anchor xml:id="endDel"/>
```

The *change* attribute used here to indicate the sequencing of these various interventions is discussed below, in section *12.7. Identifying Changes and Revisions*. The elements <addSpan> and <delSpan> are discussed in section *12.3.1.4. Additions and Deletions*.

The <metamark> element may also be used to encode the symbols etc. often found in marked-up proofs such as the following, taken from the Walt Whitman archive:

In this example, the whole of what was originally the 14th section of the poem has been marked for deletion, both by horizontal and vertical lines, and by the proofreading mark resembling the 'deleatur' or 'dele' deletion symbol to

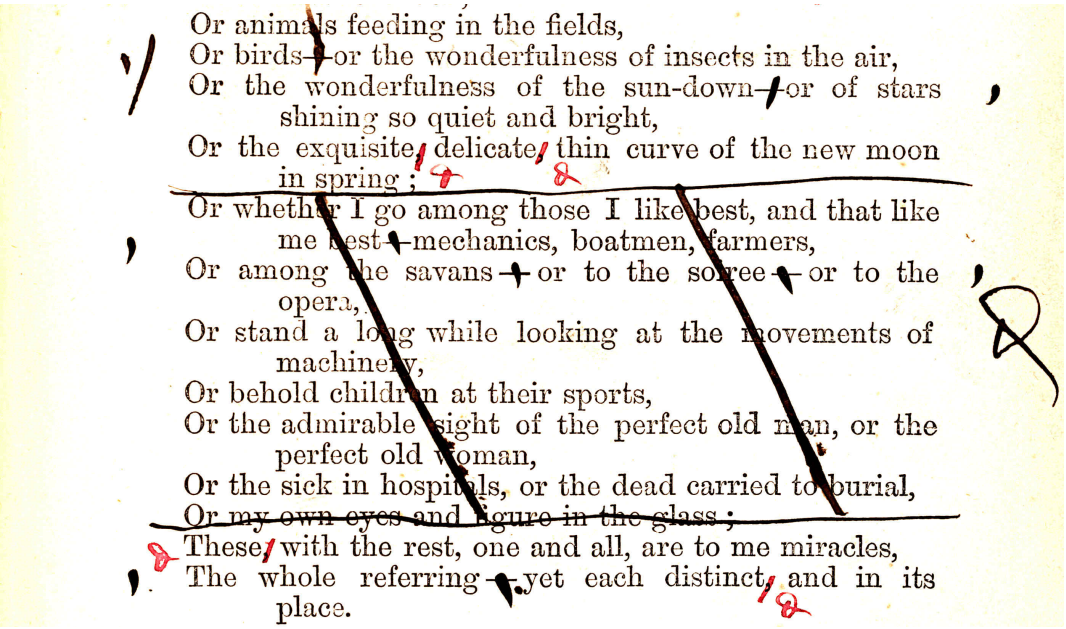



Figure 24: From a corrected proof of *Miracles* (Walt Whitman Archive)

left and right of the section. The deletion itself might be encoded by using the normal <del> or <delSpan> element, and the metamarks by the <metamark> element. This is quite a different case from that of the next example, in which the writer does not intend to suppress the content, but only to mark that it has been copied to another manuscript or reused.

This page contains internal deletions, additions, and retracings but these are semantically quite different from the apparent 'deletion' signalled by the larger of the two single vertical lines, which shows that the written material has been transferred or re-used, not deleted.

```
<surface>
  <metamark  function="used"  rend="line"
    target="#X2"/>
  <zone  xml:id="X2">
    <line>I am that halfgrown <add>angry</add> boy, fallen asleep</line>
    <line>The tears of foolish passion yet undried</line>
    <line>upon my cheeks.</line>
<!-- ... -->
    <line>I pass through <add>the</add> travels and <del>fortunes</del> of
    <retrace>thirty</retrace>
    </line>
    <line>years and become old,</line>
    <line>Each in its due order comes and goes,</line>
    <line>And thus a message for me comes.</line>
    <line>The</line>
  </zone>
  <metamark  function="used"  target="#X2">Entered - Yes</metamark>
</surface>
```

In this embedded transcription example, we class as metamarks both the long vertical line and the annotation 'Entered - yes'. Both metamarks are assumed to indicate that the whole of the written zone with identifier X2 is marked as having been used. <metamark> can be similarly used within <text> to encode the same phenomenon as part of a transcription that privileges logical over physical and layout structures.

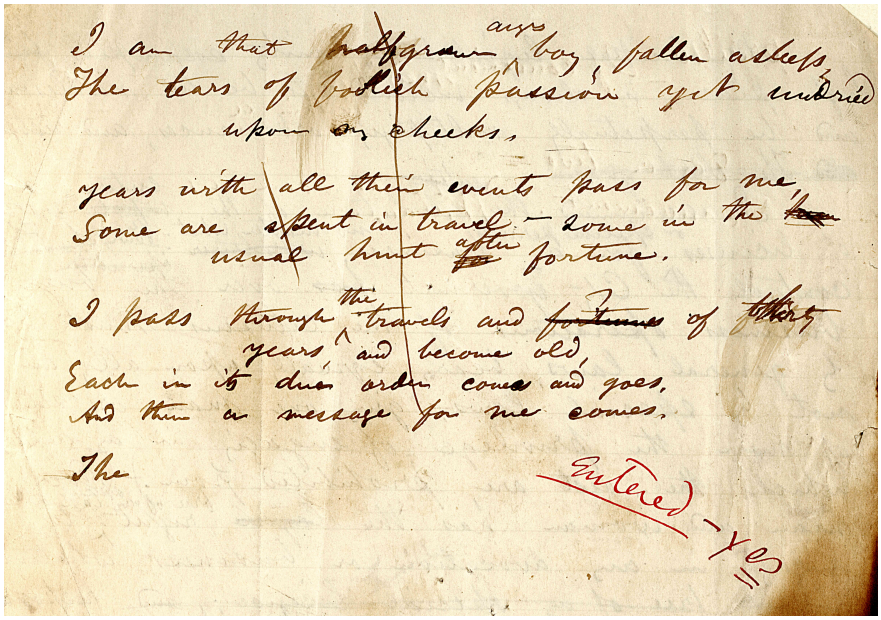



Figure 25: From "I am that halfgrown angry boy" (MS q 25), David M. Rubenstein Rare Book & Manuscript Library, Duke University.

#### 12.3.4.3 Fixation and Clarification

A writer may sometimes rewrite material a second time without significant change and in the same place. We consider this a distinct activity from addition as usually defined because no new textual material results; instead the status of existing material is reaffirmed. We may distinguish two variants of this: *fixation* where the first version was a tentative draft which is subsequently reaffirmed, for example by inking it over; and *clarification*, where the first version was badly written and has been rewritten for clarity. The element <retrace> is provided for both cases; its *cause* attribute may be used to distinguish these or other cases.

`<retrace>` contains a sequence of writing which has been retraced, for example by over-inking, to clarify or fix it.

In this simple example, taken from the papers of Henrik Ibsen, the writer wrote the word *skuldren* hastily, and then returned to it to make the letter *l* larger and clearer: We might transcribe this word as follows:

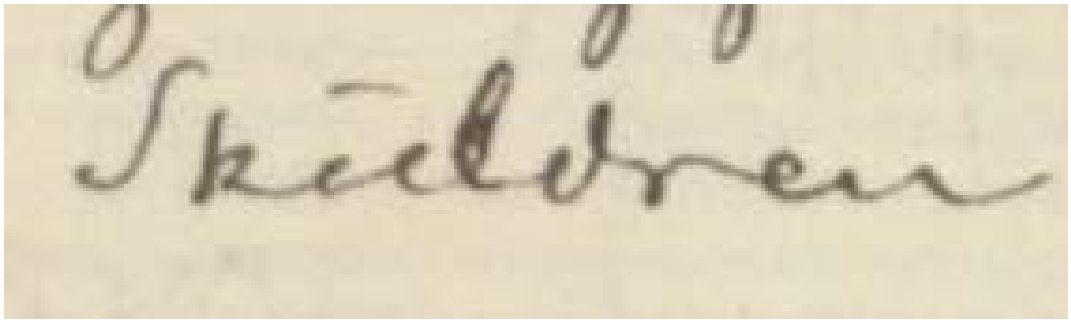



Figure 26: From autograph ms of *Peer Gynt*, Collin 2869, 4°, I.1.1, the Royal Library of Copenhagen

```
Sku<retrace cause="unclear">l</retrace>dren
```

A single rewrite may not be sufficient, and it may be that the document becomes almost unreadable as a result of repeated clarification. In the following example, we can distinguish at least three attempts to write the letters *er* in the word *bægerklang*: We might encode this by nesting the <retrace> element as follows:

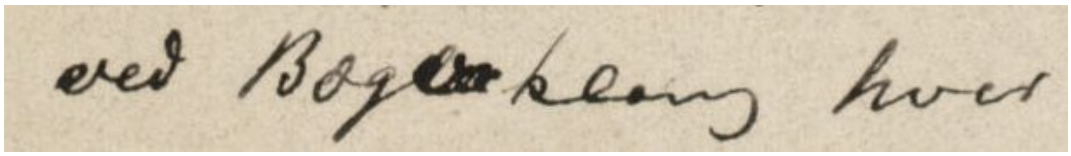


Figure 27: Detail from autograph ms *Brand* in The Royal Library, Denmark (KBK Collin 262)

```
 ved Bæg<retrace  cause="unclear"  change="#stage2">
  <retrace  cause="unclear"  change="#stage1">er</retrace>
</retrace> ...
```

The *change* attribute used here is discussed further below (*12.7. Identifying Changes and Revisions*).

The <retrace> element is used only for cases where text has been written multiple times. When metamarks and other markup-like strokes have been rewritten multiple times, the <redo> element described in the next section should be used in preference.

#### 12.3.4.4 Confirmation, Cancellation, and Reinstatement of Modifications

A writer may indicate that an alteration is itself to be altered: for example, a struck-through passage may be restored via a dotted underlining, or the underlining of a passage may be deleted by a wavy line.

The following elements are provided to represent these situations:

`<redo>` indicates one or more cancelled interventions in a document which have subsequently been marked as reaffirmed or repeated.

- `@target` points to one or more elements representing the interventions which are being reasserted.

`<undo>` indicates one or more marked-up interventions in a document which have subsequently been marked for cancellation.

- `@target` points to one or more elements representing the interventions which are to be reverted or undone.

The element <restore> (*12.3.1.6. Cancellation of Deletions and Other Markings*) is provided for the comparatively simple case where a simple deletion is marked as having been subsequently cancelled. The <undo> element discussed here is more widely applicable and may be used for any kind of cancellation. It points to the element or elements which are being cancelled. These components need not be contiguous, provided that the cancellation is clearly a single act; each distinct act of cancellation requires a distinct <undo> element, however. Either of the attributes *target* or *spanTo* may be used to indicate the passages concerned.

Consider the following imaginary example: We hypothesize that the text has gone through three states or

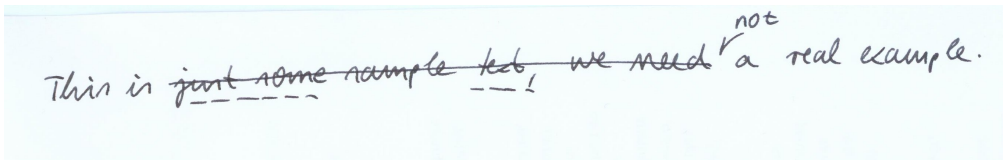



Figure 28: Imaginary example demonstrating restoring and undoing

changes, as follows:

1. This is just some sample text, we need a real example.

2. This is not a real example.
3. This is just some text, not a real example.

This sequence of events might be encoded as follows:

```
 This is <del  change="#s2"  rend="overstrike">
  <undo  spanTo="#Xa"  rend="dotted"
    change="#s3"/>just some <anchor  xml:id="Xa"/> sample <undo  spanTo="#Xb"  rend="dotted"
    change="#s3"/>text, <anchor  xml:id="Xb"/> we need
</del>
<add  change="#s2">not</add> a real example.
```

using two <undo> elements, each with a *spanTo* attribute, to delimit the two parts of the deletion which were reverted at change s3. Note that in this case, since *target* is not supplied, it is the effect of the parent element (the <del>) which is assumed to be undone.

Alternatively, we might more economically use the generic <seg> element to delimit the two sequences whose deletion is being reverted, and then use the *target* attribute on a single <undo> element:

```
 This is <del  change="#s2"  rend="overstrike">
  <seg  xml:id="X-a">just some</seg> sample <seg  xml:id="X-b">text</seg>, we
need
</del>
<add  change="#s2">not</add> a real example.

<undo  target="#X-a #X-b"  rend="dotted"
  change="#s3"/>
```

#### 12.3.4.5 Transpositions

A *transposition* occurs when metamarks are found in a document indicating that passages should be moved to a different position. Typically this may be done using arrows, asterisks or numbers, or other means. By definition the result of a transposition is not present in the document, and should not therefore be encoded, if the intention is to represent the actual appearance of the document. Instead, the following elements may be used to indicate the intended reordering:

`<listTranspose>` supplies a list of transpositions, each of which is indicated at some point in a document typically by means of metamarks.

`<transpose>` describes a single textual transposition as an ordered list of at least two pointers specifying the order in which the elements indicated should be re-combined.

Consider for example, the following extract from an Ibsen manuscript The underlined numbers 1 and 2 here

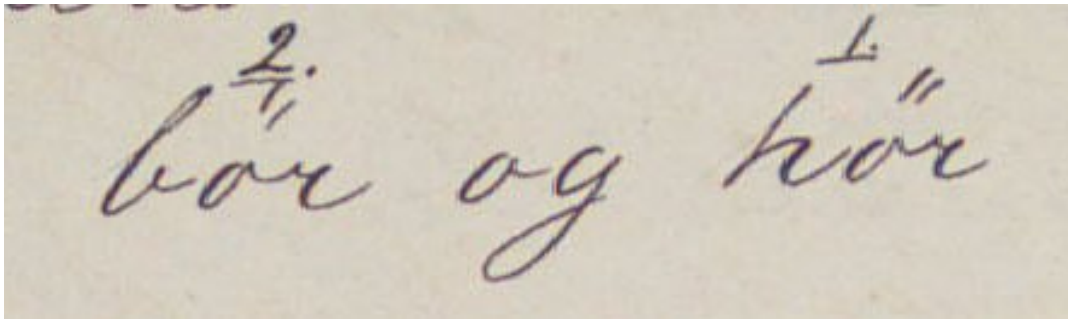



Figure 29: Extract from autograph *Digte* (Poems) NBO Ms. 4º 1110a

indicate that, although the word *bör* precedes the word *hör* in the text, the order of the two words should be reversed. We may encode this as follows:

```
<seg  xml:id="ib01">bör</seg>
<metamark  rend="underline"
  function="transposition"  target="#ib01"  place="above">2.</metamark> og
<seg  xml:id="ib02">hör</seg>
<metamark  rend="underline"
  function="transposition"  target="#ib02"  place="above">1.</metamark>
<!-- either nearby or elsewhere, typically inside <profileDesc> or <standOff>: -->
<listTranspose>
  <transpose>
    <ptr  target="#ib02"/>
    <ptr  target="#ib01"/>
  </transpose>
</listTranspose>
```

Note the use of the generic <seg> element to identify the sections of text being transposed. The following example uses an embedded transcription approach, which typically identifies lines of writing with the <line> element. This makes it trivial to refer to the transposed line, but when encoding transposition within <text> the encoder will need to find a way to identify the line with another element, such as <seg>.

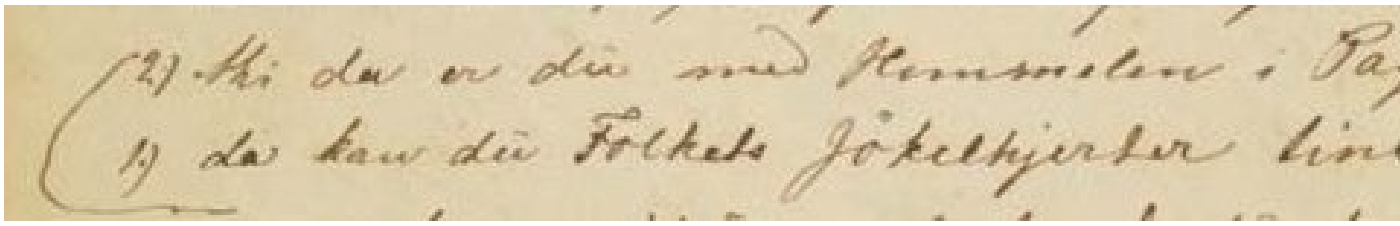


Figure 30: Detail from autograph ms of *Den episke Brand* (KBK Collin 2869)

```
<line  xml:id="ib3">
  <metamark  function="transposition"
    place="margin-left">2.)</metamark> thi da er du med Himmelen i Pagt; —
</line>
<line  xml:id="ib4">
  <metamark  function="transposition"
    place="margin-left">1.)</metamark> da kan du
Folkets Jøkelhjerter tine;
</line>
<!-- either nearby or elsewhere, typically inside <profileDesc> or <standOff>: -->
<listTranspose>
  <transpose>
    <ptr  target="#ib4"/>
    <ptr  target="#ib3"/>
  </transpose>
</listTranspose>
```

When transposition is made, the whole element indicated is understood to be moved, not just its contents. In the above example, the metamarks are thus understood to be moved along with the lines to which they apply.

One or more <listTranspose> elements may be supplied either embedded within the text, in the <profileDesc> of the header, or in a <standOff> depending on local preference. Each <listTranspose> can contain one or more <transpose> elements, each of which defines a single transposition.

#### 12.3.4.6 Alternative Readings

Figure 31: Detail from autograph manuscript of the second version of *Lalla Rookh*, Pierpont Morgan MA 310

In this example two alternative readings are provided, but no preference is indicated. While the author apparently first composed the line 'Alone before his native river -', at some later point, he entertained the possibility of using the word *beside* instead of *before*. The manuscript supplies no indication of which word Moore favours at this point, although in fact, in the first printed edition of *Lalla Rookh* the word *beside* was chosen.

The element <alt> provided by the `linking` module gives a simple way of encoding the state of this manuscript, as follows:

```
Alone <seg  xml:id="alt1">before</seg>
<add  place="above"  xml:id="alt2">beside</add> his native river —

<alt  target="#alt1 #alt2"  mode="excl"
  weights="0 1"/>
```

The *mode* attribute here indicates that the two possible readings indicated by the *target* attribute are mutually exclusive. The *weights* attribute indicates the relative importance or preference to be attached to the two readings on a scale running from zero (most improbable) to one (most probable). In this case, we have a very strong preference for the second reading because this is the one that appears in the published version of the poem. The <alt> element is further discussed in section *17.8. Alternation*.

#### 12.3.4.7 Instant Corrections

The use of elements such as <del> and <add> necessarily implies that the modifications they indicate were made at some time after the original writing. An exception to this is where a false start or 'instant' correction has been identified: the author starts to write, and then immediately corrects what has been written.

The *instant* attribute defined by this module may be used on any element which is a member of the `att.editLike` class to modify this default assumption. When the value of *instant* is set to true, the addition or deletion is considered to belong to the same change as its parent element, while `false` means some change later than that of its parent.

An example of false start or instant correction can be seen in the following line: in which we can detect the

Figure 32: Detail fron *[I am a curse]*, one of the drafts for Whitman's *Song of Myself*

following sequence of events:

1. The letter T is written and then immediately deleted
2. The word The is written, deleted, and replaced by the word His
3. The added word His is then deleted
4. The initial letter i of the words *iron necklace* is overwritten with a capital I

To indicate that the first of these acts must have taken place during the main act of writing, before the other deletion and additions, we might encode this revision campaign as follows:

```
<del instant="true">T</del>
<mod type="subst">
  <del>The</del>
  <add place="above">
    <del rend="overstrike">His</del>
  </add>
</mod>
<mod type="subst">
  <del rend="overwritten">i</del>
  <add place="superimposed">I</add>
</mod>ron necklace
```

## 12.4 Aspects of Layout

The following methods are available to capture general aspects of the layout of material on a page where this is considered important. Within the <sourceDoc> element, as already indicated, the element <surface> and <surfaceGrp> enable the encoder to represent directly the structure of a codex as gatherings or quires, leaves, and surfaces, as in the following example:

```
<sourceDoc>
  <surfaceGrp type="quire" n="1">
    <surfaceGrp type="leaf" n="1">
      <surface type="recto" n="1r">
<!-- ... -->
      </surface>
      <surface type="verso" n="1v">
<!-- ... -->
      </surface>
    </surfaceGrp>
    <surfaceGrp type="leaf" n="2">
      <surface type="recto" n="2r">
<!-- ... -->
      </surface>
      <surface type="verso" n="2v">
<!-- ... -->
      </surface>
    </surfaceGrp>
<!-- other leaves in first quire -->
  </surfaceGrp>
<!-- other quires here -->
</sourceDoc>
```

In some cases, it may be preferable to define <surface>s corresponding with each two page opening, for example where it is clear that the writer regarded each such opening as a single writing surface, with written zones or other features crossing the page divide. An example is shown here:

The coloured lines added to this image indicate a number of zones of writing, colour coded to indicate the order in which they were written (purple, then green, then red). For example, the zone marked in red on the left contains a note referring to the purple zone on the right.

This approach assumes that the transcription will primarily be organized in the same way as the physical layout of the source, using embedded transcription elements. Alternatively, where a non-embedded transcription has been provided, using the <text> element, it is still possible to record gathering beginnings, page beginnings, column beginnings, line beginnings etc. in the source, using the elements described in section *3.11. Reference Systems*. Detailed metadata about the physical make-up of a source will usually be summarized by the <physDesc> component of an <msDesc> element discussed in *11.7. Physical Description*.

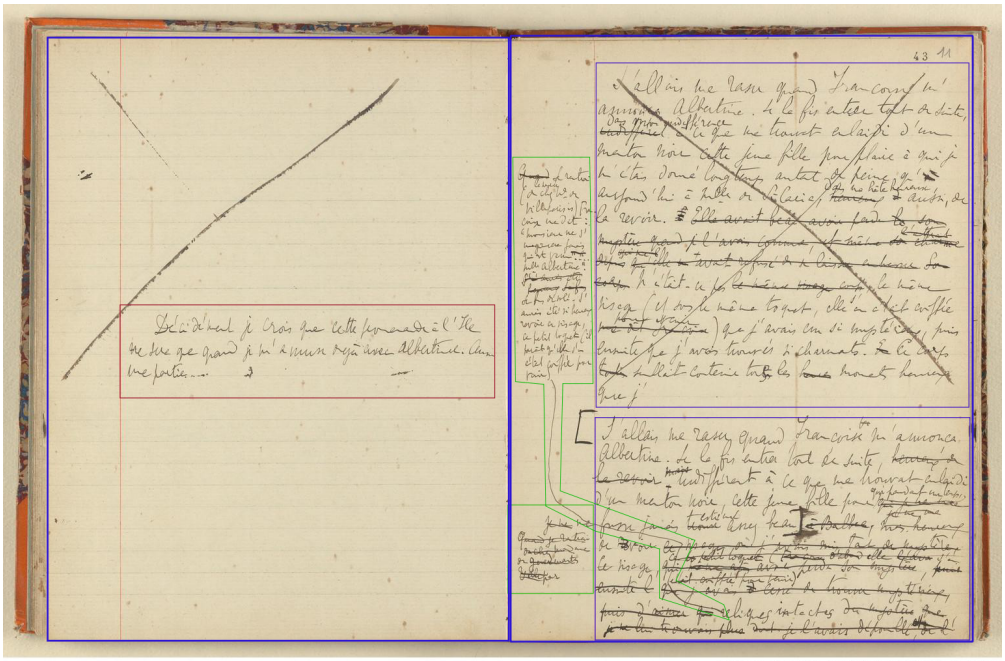



Figure 33: Opening from autograph ms of Proust's *A la recherche du temps perdu* (f 42v-43r)

### 12.4.1 Space

The author or scribe may have left space for a word, or for an initial capital, and for some reason the word or capital was never supplied and the space left empty. The presence of significant space in the text being transcribed may be indicated by the <space> element.

`<space>` (space) indicates the location of a significant space in the text.

`@resp` (responsible party) (responsible party) indicates the individual responsible for identifying and measuring the space.

Note that this element should not be used to mark normal inter-word space or the like.

In line 694 of Chaucer's *Wife of Bath's Prologue* in the Holkham manuscript the scribe has left a space for a word where other manuscripts read *preestes*:

```
By
god if wommen had writen storyes As <space  quantity="7"  unit="chars"/> han within her
oratoryes
```

The <supplied> element discussed in the previous section may be used to supply the text presumed missing:

```
By god if wommen had writen storyes As
<supplied  reason="space"  resp="#ETD"
  source="#Hg">preestes</supplied> han within
her oratoryes
```

Here, the fact of the space within the manuscript is indicated by the value of the *reason* attribute. The source of the supplied text is shown by the value of the *source* attribute as the Hengwrt manuscript; the transcriber responsible for supplying the text is ES.

### 12.4.2 Lines

One of the more common forms of modification encountered in written documents of any kind is the presence of lines written under, beside, or through the text. Such lines may be of various types: they may be solid, dashed or dotted, doubled or tripled, wavy or straight, or a combination of these and other renderings. The line may be used for emphasis, or to mark a foreign or technical term, or to signal a quotation or a title, etc.: the elements <emph>, <foreign>, <term>, <mentioned>, or <title> may be used for these. Where the line has a clear paratextual function the <metamark> element may be considered more appropriate. Frequently, a scholar may judge that a line is used to delete text: the <del> element is available to indicate this. In all these cases, the *rend* attribute may be used to supply further details concerning the style of the line. Thus, Lawrence's deletion by strike-through of *my* in the autograph of *Eloi, Eloi, lama sabachthani* may be encoded:

```
For I hate this <del  rend="strikethrough"  hand="#dhl">my</del> body, which is so
dear to me
```

There will be instances, however, where a scholar wishes only to register the occurrence of lines in the text, without making any judgement as to what the lines signify. In these cases the <hi> element may be used, with the *rend* attribute to mark the style of line. In the manuscript of a letter by Robert Browning to George Moulton-Barrett the underlining of the phrase *had obtained all the letters to Mr Boyd* may be marked-up as follows:

```
I have once — by
declaring I would prosecute by law — hindered a man's proceedings who
<hi  rend="underline">had obtained all the letters to Mr Boyd</hi>
```

The above examples presume the common case where a single word or phrase is marked by a line, with no doubt as to where the marking begins or ends and with no overlapping of the area of text with other marked areas of text. Where there is doubt, the <certainty> element may be used to record the doubt. In the Browning example cited above the underlining actually begins half-way under *who*, and this uncertainty could be remarked as follows:

```
I have once — by declaring I would prosecute
by law — hindered a man's proceedings who <hi  xml:id="cstart1"  rend="underline">had
obtained all the letters to Mr Boyd</hi>
<!-- ... -->
<certainty  target="#cstart1"  locus="start"
  degree="0.70">
  <desc>may begin with previous word</desc>
</certainty>
```

Where the area of text marked overlaps other areas of text, for example crossing a structural division, one of the spanning mechanisms mentioned above must be used; for example where the line is thought to mark a deletion, the <delSpan> element may be used. Where it is desired simply to record the marking of a span of text in circumstances where it is not possible to surround the text with a <hi> element, the  element may be used with the *rend* or *type* attribute indicating the style of line-marking.

More work needs to be done on clarifying the treatment of other textual features marked by lines which might so overlap or nest. For example, in many Middle English manuscripts (e.g. the Jesus and Digby verse collections), marginal sidebars may indicate metrical structure: couplets may be linked in pairs, with the pairs themselves linked into stanzas. Or, marginal sidebars may indicate emphasis, or may point out a region of text on which there is some annotation: in many manuscripts of Chaucer's *Wife of Bath's Prologue* lines 655–8 are marked with nesting parentheses against which the scribe has written *nota*.

Such features could be captured by use of the <note> element, containing a prose description of the manuscript at this point, enhanced by a link to a visual representation (or facsimile) of the feature in question. For example, in the Chaucer example just cited, one may wish to record that the *nota* is written in the Hengwrt manuscript in the right margin against a single large left parenthesis bracketing the four lines, with two right parentheses in the right

margin bracketing two overlapping pairs of lines: the first and third, the second and fourth. The <note> element allows us to record that the scribe wrote *nota*, but is not well-adapted to show that the *nota* points both at all four lines and at two pairs of lines within the four lines. The <metamark> element discussed in section *12.3.4.2. Metamarks* above provides better facilities for this kind of complex annotation.

## 12.5 Transcription and Ruby

These Guidelines also provide special elements to support the encoding of ruby annotations, which are common in East Asian textual traditions. These elements provide a method of capturing a specific type of annotation, in addition to the generic methods like the <note> or <interp> elements. Both the specific and general methods should integrate well with the transcriptional elements described above, allowing authorial and scribal features to be captured in conjunction with ruby base text and annotations. See *3.4.2. Ruby Annotations* for more information about these elements.

## 12.6 Headers, Footers, and Similar Matter

Such information as page numbers, signatures, or catchwords may be recorded in a specialized <fw> element provided for that purpose. Although the name derives from the term *forme work*, used in description of early printed documents (the 'forme' being the block used to hold movable type), the <fw> element may be used for such features of any document, written or printed. Note that the purpose of this element is to record page numbers etc. *actually present* in the document being encoded, not necessarily to provide a complete or accurate pagination of it.

Information about pagination etc. may also be provided using the *n* attribute of the <pb> or <gb> elements, or by other appropriate <milestone> elements, as further discussed in section *3.11. Reference Systems*: since this information is usually provided by the encoder, it is not subject to the constraint that it should be present only if textually present in the source being encoded. In text-critical situations it may be useful to provide both a normalized version of the pagination and a representation of the catch-word or numbering, especially when the latter presents a variant reading, or is significant for compositor identification.

`<fw>` (forme work) contains a running head (e.g. a header, footer), catchword, or similar material appearing on the current page.

The <fw> element may be used to encode any of the unchanging portions of a page forme, such as:

- running heads (whether repeated or changing on every page, or alternating pages)
- running footers
- page numbers
- catch-words
- other material repeated from page to page, which falls outside the stream of the text

It should not be used for marginal glosses, annotations, or textual variants, which should be tagged using <gloss>, <note>, or the text-critical tags described in chapter *13. Critical Apparatus*, respectively.

For example:

```
<fw type="head" place="top-centre">Poëms.</fw>
<fw type="pageNum" place="top-right">29</fw>
<fw type="sig" place="bot-centre">E3</fw>
<fw type="catch" place="bot-right">TEMPLE</fw>
```

## 12.7 Identifying Changes and Revisions

A major purpose of genetic editing is the identification of 'revision campaigns' or, more generally, *changes*. An editor may wish to regard a particular set of alterations (deletions, additions, substitutions, transpositions, etc.) or any other act of writing as a single object for which we use the general term *change*, to indicate both that one or more of such phenomena preceded or followed another and also to indicate that they are related in some way, for example that one is a consequence of the other. They might also wish to group together certain revisions, regardless of when they might have occurred, based on a variety of other shared characteristics (e.g., corrections of factual errors or revisions that incorporate suggestions made by a given reader). To document this we need:

- a system to assign phenomena to a particular 'change' as defined above
- a way to characterize each such change, in itself and in relation to others.

The element <creation> (within the TEI header profile description) contains all information relating to the genesis or production of a text. It may contain a <listChange> element which contains a number of <change> elements, one for each set of alterations identified:

`<listChange>` groups a number of change descriptions associated with either the creation of a source text or the revision of an encoded text.

`@ordered` indicates whether the ordering of its child <change> elements is to be considered significant or not.

`<change>` (change) documents a change or set of changes made during the production of a source document, or during the revision of an electronic file.

In the following example an editor has identified four distinct sets of alterations:

```
<profileDesc>
  <creation>
    <listChange  ordered="true">
      <change  xml:id="ST-1">First stage, written in ink </change>
      <change  xml:id="ST-2">Second stage, with revisions written in the author's hand
            using pencil</change>
      <change  xml:id="ST-3">Fixation of the pencilled revisions together with further
            revisions in the author's hand using ink</change>
      <change  xml:id="ST-4">Additions in a different hand, probably at a later
            stage</change>
    </listChange>
  </creation>
</profileDesc>
```

The <listChange> element carries an attribute *ordered*, which can take the values `true` or `false` (the default). The attribute specifies whether the order of child elements signifies a temporal order for the revision campaigns which they document. In the example above, the editor has asserted that the four sets distinguished are ordered chronologically according to the order of the <change> elements. If necessary, <listChange> elements can be nested hierarchically. This may be helpful in two cases. Firstly one can build up hypotheses about related revisions step-by-step, starting with <change> elements of smaller coverage, whose members are certainly related, and then in a subsequent pass grouping these in turn, thereby extending their reach.

```
<profileDesc>
  <creation>
    <listChange>
      <change  xml:id="o">An unrelated change note</change>
      <listChange  xml:id="m">
        <change  xml:id="m1">Alterations on one manuscript page, certainly
              related</change>
        <change  xml:id="m2">Alterations on another manuscript page, certainly
              related</change>
      </listChange>
      <change  xml:id="p">Another unrelated change note</change>
    </listChange>
  </creation>
</profileDesc>
```

A nested <listChange> element is also useful to indicate a *partial* ordering of <change> elements.

```
<listChange ordered="true">
  <change xml:id="ST1">The first stage</change>
  <listChange ordered="false">
<!-- We have no information about the order of these changes, except
that they both followed ST1 and preceded STX -->
    <change xml:id="ST-rev1">A revision of the first stage</change>
    <change xml:id="ST-rev2">Another revision of the first stage</change>
  </listChange>
  <change xml:id="STX">The last stage</change>
</listChange>
```

In addition to the possibility of being ordered by their sequence within a <listChange> element, <change> elements may carry a number of attributes from the `att.datable` class (*period*, *when*, *notBefore*, *notAfter*, *from*, and *to*) which allow each element to be dated as exactly or inexactly as necessary, in the same way as is currently possible for the TEI <date> element.

```
<profileDesc>
  <creation>
    <date notAfter="1816-07-18"/>
    <listChange ordered="true">
      <change xml:id="mod1" when="1816-07-16">The first draft of
      <title>Persuasion</title>, completed by the date <date>July 16 1816</date>
            which is written after the word <q>Finis</q> at <ref target="#pers-30">page
               30</ref>.</change>
      <change xml:id="mod2"
        notBefore="1816-07-16">After the <date>16th of July</date>
            Austen starts revision of the two final chapters, by rewriting the end and
            adding a new zone (<ref target="#transp-1">pages 32-35</ref>) to be
inserted at
      <ref target="#insertion-p1">page 19</ref>. This stage is documented by the
            deletion of the date (<date>July 16 1816</date>) at <ref target="#pers-30">page
               30</ref>, and the addition of more text and of a new date (<date>July
18.
               1816</date>) at <ref target="#pers-31">page 31</ref>
      </change>
      <change notBefore="1816-07-18">Before publication, after <date>July 18th,
               1816</date> chapters 10-11 were broken into three chapters, 10, 11, 12,
as
            witnessed by the print.</change>
    </listChange>
  </creation>
</profileDesc>
```

Each <change> element, apart from declaring a distinct moment or phase in the creation of the document, may also contain references to other annotations contained within the <teiHeader> or in the document (as shown in the previous example). Such references, along with the textual content, are purely documentary. The association between a textual component and a <change> element is always made explicitly, either by using the *target* attribute on the <change> element to point to one or more textual elements, or by pointing from the element or elements concerned to the <change> element by means of their *change* attribute. If a <change> element is associated with some element, it is also associated with all of that element's children, unless otherwise indicated, for example by a new value for the *change* attribute.

In the following simple example, the text at one stage read 'This is a mouse', and at the next 'This is a house mouse':

```
<line change="#firstStage">This is a <add change="#secondStage">house</add>
mouse.</line>
```

In this example, however, the text originally read 'This is a house', and subsequently 'This is a mouse':

```
<line change="#firstStage">This is a <mod type="subst" change="#secondStage">
   <del>house</del>
   <add>mouse</add>
  </mod>.</line>
```

Note that in this case both the deletion and the addition are associated with the second <change> element. The word 'house', because its deletion forms part of this second set of alterations, must have been present originally, whereas the word 'mouse' must have been added during the second set of alterations.

Elements such as <add> and <del> and the like carry an implied semantics concerning the order in which events in the writing of a document was carried out: something which is deleted must have been written before it was deleted; something which is added must have been added at a later stage of the writing. Even when a combination of such elements is used, the chronology can usually be inferred (see further *12.3.3.2. Use of the gap, del, damage, unclear, and supplied Elements in Combination*). Explicit indication of the set of alterations to which some modification belongs is mostly useful in situations where all the alterations identified in a document are to be grouped, for example chronologically.

The interpretation of <change> elements with respect to a particular text passage is based on a number of implicit assumptions and constraints which have the effect of minimizing the amount of tagging necessary. The system is also flexible enough to support an explicit distinction between acts of writing and textual alterations, since either of these can be associated with changes described in the encoding. The following example shows an encoding in which the same passage is transcribed twice, once from a documentary perspective, and once from a textual one:

```
<profileDesc>
  <creation>
    <listChange ordered="true">
      <change target="#zone_1 #subst_3">First stage, written in ink by a
            scribe</change>
      <change target="#zone_2 #mod_1 #line_1 #line_2 #subst_1 #subst_2 #subst_4
#delSpan_1">Revised by Goethe using pencil</change>
      <change target="#redo_1 #redo_2 #redo_3 #subst_1 #subst_2 #delSpan_1
#add_1">Fixation of the revised passages and further revisions by Goethe using
            ink</change>
    </listChange>
  </creation>
</profileDesc>
<sourceDoc>
  <surface>
    <zone xml:id="zone_1">
      <line xml:id="line_1">
        <handShift new="#g_bl"/>
        <retrace hand="#g_t" xml:id="redo_1">Nun</retrace>
      </line>
      <line>
        <handShift new="#jo_t"/>Ihr wanſtige Schuften mit den Feuerbacken</line>
      <line xml:id="line_2">
        <handShift new="#g_bl"/>
        <retrace hand="#g_t" xml:id="redo_2">feiſt</retrace>
      </line>
      <line>Ihr glüht ſo recht vom Höllen Schwefel ſatt.</line> [...] </zone>
```

```
  </surface>
</sourceDoc>
<text>
  <body>
    <l  n="11656">
      <subst  xml:id="subst_1">
        <del>Ihr</del>
        <add>Nun</add>
      </subst> wanftige Schuften mit den Feuerbacken</l>
    <l  n="11657">Ihr glüht ſo recht vom Höllen Schwefel <subst  xml:id="subst_2">
        <del>ſatt</del>
        <add>feiſt</add>
      </subst>.</l>
  </body>
</text>
```

The documentary transcription stresses the writing process, while the textual transcription emphasizes textual alterations. In either case, the change of writing activity associated with a particular feature in the transcript is explicitly indicated. From the documentary perspective, by assigning particular modifications to a specific <change> element, we describe the writing process, in that they specify which segment has been written when . From the textual perspective, the markup concentrates simply on the existence of textual alterations and makes no explicit claims about the order of writing.

## 12.8 Other Primary Source Features not Covered in these Guidelines

We repeat the advice given at the beginning of this chapter, that these recommendations are not intended to meet every transcriptional circumstance ever likely to be faced by any scholar. They are intended rather as a base to enable encoding of the most common phenomena found in the course of scholarly transcription of primary source materials. These guidelines particularly do not address the encoding of physical description of textual witnesses: the materials of the carrier, the medium of the inscribing implement, the organisation of the carrier materials themselves (as quiring, collation, etc.), and authorial instructions or scribal markup, except insofar as these are involved in the broader question of manuscript description, as addressed by the `msdescription` module described in chapter *11. Manuscript Description*.

## 12.9 Module for Transcription of Primary Sources

The module described in this chapter makes available the following components:

**Module transcr: Transcription of primary sources**

- *Elements defined*: addSpan am damage damageSpan delSpan ex facsimile fw handNotes handShift line listTranspose metamark mod path redo restore retrace secl sourceDoc space subst substJoin supplied surface surfaceGrp surplus transpose undo zone
- *Classes defined*: att.coordinated att.global.change att.global.facs

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 13 Critical Apparatus

Scholarly editions of texts, especially texts of great antiquity or importance, often record some or all of the known variations among different *witnesses* to the text. Witnesses to a text may include authorial or other manuscripts, printed editions of the work, early translations, or quotations of a work in other texts. Information concerning variant readings of a text may be accumulated in highly structured form in a critical apparatus of variants. This chapter defines a module for use in encoding such an apparatus of variants, which may be used in conjunction with any of the modules defined in these Guidelines. It also defines an element class which provides extra attributes for some elements of the core tag set when this module is selected. In printed critical editions, the apparatus takes the form of highly-compressed notes at the bottom of each page. TEI's critical apparatus module allows variation to be encoded so that such notes may be generated, but it also models the variation so that, for example, interactive editions in which readers can choose which witness readings to display are possible.

Information about variant readings (whether or not represented by a critical apparatus in the source text) may be recorded in a series of *apparatus entries*, each entry documenting one *variation*, or set of readings, in the text. Elements for the apparatus entry and readings, and for the documentation of the witnesses whose readings are included in the apparatus, are described in section *13.1. The Apparatus Entry, Readings, and Witnesses*. Special tags for fragmentary witnesses are described in section *13.1.5. Fragmentary Witnesses*. The available methods for embedding the apparatus in the rest of the text, or for linking an external apparatus to the text of the edition, are described in section *13.2. Linking the Apparatus to the Text*. Finally, several extra attributes for some tags of the core tag set, made available when the additional tag set for text criticism is selected, are documented in section *12.3.1.1. Core Elements for Transcriptional Work*.

Scholarly practice in representing critical editions differs widely across disciplines, time periods, and languages. The TEI does not make any recommendations as to which text-critical methods are best suited to any given text. Editors will wish to consider questions such as:

- What source documents will be used? Are there many witnesses, few, or one? Are the sources relatively close copies or not?
- Will there be a single 'base' text? Or will witnesses be separately transcribed?
- If a single base text will be used, will it be that of a particular witness, or will the editor attempt to reconstruct an ideal or original text?
- Will each reading in an apparatus entry record every attestation (a 'positive' apparatus), or merely witnesses that deviate from the base text (a 'negative' apparatus)?
- Will the readings of most or all witnesses be represented in the apparatus, or only a selection the editor deems relevant?
- What level of variation will require distinguishing one witness reading from another? For example, will the editor consider an abbreviated word in a witness as agreeing with the base text, or not?
- Will conjectures (variant readings suggested by an editor) be treated differently than readings found in witnesses?
- Will there be a need to distinguish different types of variation, for example orthographic vs. morphological or lexical variants?

Different editorial methodologies will produce different answers to these questions, and those answers may influence choices of markup used in the edition. For example, if there will be multiple witness transcriptions and a single apparatus, then the double end-point attachment method may be the best choice of apparatus linking style. The parallel segmentation method may present several advantages to editors producing an edition with a single base text. Editors of single-source editions may care to note material aspects of the text (such as <damage> or <unclear> text). On the other hand, editors attempting to synthesize an ideal or original text from many witnesses may feel little need to represent the material aspects of individual witnesses. Editors wishing to distinguish witness readings from conjectures by modern editors may wish to use *wit* to indicate the former and *source* for the latter. Differences in types of variation might be marked using *type* or *ana* on the <rdg> element.

Many examples given in this chapter refer to the following texts of the opening (usually just line 1) of Chaucer's *Wife of Bath's Prologue*, as it appears in each of the four different manuscripts

- Ellesmere, Huntingdon Library 26.C.9 (**El**)
- Hengwrt, National Library of Wales, Aberystwyth, Peniarth 392D (**Hg**)

- British Library Lansdowne 851 (**La**)
- Bodleian Library Rawlinson Poetic 149 (**Ra2**)

## 13.1 The Apparatus Entry, Readings, and Witnesses

This section introduces the fundamental markup methods used to encode textual variations:

- the <app> element for entries in the critical apparatus: see section *13.1.1. The Apparatus Entry*.
- elements for identifying individual readings: see section *13.1.2. Readings*.
- ways of grouping readings together: see section *13.1.3. Indicating Subvariation in Apparatus Entries*.
- methods of identifying which witnesses support a particular reading, and for describing the witnesses included in the apparatus: see section *13.1.4. Witness Information*.
- elements for indicating which portions of a text are covered by fragmentary witnesses: see section *13.1.5. Fragmentary Witnesses*.

The <app> element is in one sense a more sophisticated and complex version of the <choice> element introduced in *3.5.1. Apparent Errors* as a way of marking points where the encoding of a passage in a single source may be carried out in more than one way. Unlike <choice>, however, the <app> element allows for the representation of many different versions of the same passage taken from different sources.

### 13.1.1 The Apparatus Entry

Individual textual variations are encoded using the <app> element, which groups together all the readings constituting the variation. The identification of discrete textual variations or apparatus entries is not a purely mechanical process; different editors will group readings differently. No rules are given here as to how to collect readings into apparatus entries.

The individual apparatus entry is encoded with the <app> element:

`<app>` (apparatus entry) contains one entry in a critical apparatus, with an optional lemma and usually one or more readings or notes on the relevant passage.

- `@type` classifies the variation contained in this element according to some convenient typology.
- `@from` identifies the beginning of the lemma in the base text.
- `@to` identifies the endpoint of the lemma in the base text.
- `@loc` (location) indicates the location of the variation, when the location-referenced method of apparatus markup is used.

The attributes *loc*, *from*, and *to*, are used to link the apparatus entry to the base text, if present. In such cases, several methods may be used for such linkage, each involving a slightly different usage for these attributes. Linkage between text and apparatus is described below in section *13.2. Linking the Apparatus to the Text*. For the use of the <app> element without a base text, see *13.2.3. The Parallel Segmentation Method*.

Each <app> element usually comprises one or more readings, which in turn are encoded using the <rdg> or other elements, as described in the next section. A very simple partial apparatus for the first line of the *Wife of Bath's Prologue* might take a form something like this:

```
<app>
  <rdg  wit="#El">Experience though noon Auctoritee</rdg>
  <rdg  wit="#La">Experiment thouh noon Auctoritee</rdg>
  <rdg  wit="#Ra2">Eryment though none auctorite</rdg>
</app>
```

Of course, in practice the apparatus will be somewhat more complex. Specifically, it may be desired to record more obviously that manuscripts El and La agree on the words 'noon Auctoritee', to indicate a preference for one reading, etc. The following sections on readings, subvariation, and witness information describe some of the more important complications which can arise.

### 13.1.2 Readings

Individual readings are the crucial elements in any critical apparatus of variants. The following elements should be used to tag individual readings within an apparatus entry:

`<lem>` (lemma) contains the lemma, or base text, of a textual variation.

`<rdg>` (reading) contains a single reading within a textual variation.

N.B. the term *lemma* is used here in the text-critical sense of 'the reading accepted as that of the original or of the base text'. This sense differs from that in which the word is used elsewhere in the Guidelines, for example as in the attribute *lemma* where the intended sense is 'the root form of an inflected word', or 'the heading of an entry in a reference book, especially a dictionary'.

In recording readings within an apparatus entry, the <rdg> element should always be used; each <app> usually contains at least one <rdg>, though it may contain only <note>s.

The <lem> element may also be used to record the base text of the source edition, to mark the readings of a base witness, to indicate the preference of an editor or encoder for a particular reading, or (e.g. in the case of an external apparatus) to indicate precisely to which portion of the main text the variation applies. Those who prefer to work without the notion of a base text or who are not using the parallel segmentation method may prefer not to use it at all. How it is used depends in part on the method chosen for linking the apparatus to the text; for more information, see section *13.2. Linking the Apparatus to the Text*.

Readings may be encoded individually, or grouped for clarity using the <rdgGrp> element described in section *13.1.3. Indicating Subvariation in Apparatus Entries*.

As members of the attribute class `att.textCritical`, both of these elements inherit the following attributes:

`att.textCritical` defines a set of attributes common to all elements representing variant readings in text critical work.

- `@type` classifies the reading according to some useful typology.
- `@cause` classifies the cause for the variant reading, according to any appropriate typology of possible origins.
- `@varSeq` (variant sequence) provides a number indicating the position of this reading in a sequence, when there is reason to presume a sequence to the variants.
- `@hand [att.written]` points to a <handNote> element describing the hand considered responsible for the content of the element concerned.

. <rdg> (but not <rdgGrp>) is also a member of `att.witnessed`:

`att.witnessed` provides attributes used to identify the witnesses supporting a particular reading in a critical apparatus.

- `@wit` (witness or witnesses) contains a space-delimited list of one or more pointers indicating the witnesses which attest to a given reading.

These elements also inherit the following attributes from the `att.global.responsibility` class:

`att.global.responsibility` provides attributes indicating the agent responsible for some aspect of the text, the markup or something asserted by the markup, and the degree of certainty associated with it.

- `@resp` (responsible party) indicates the agency responsible for the intervention or interpretation, for example an editor or transcriber.
- `@cert` (certainty) signifies the degree of certainty associated with the intervention or interpretation.

As elsewhere, these attributes may be used to indicate the person responsible for the editorial decision being recorded, and also the degree of certainty associated with that decision by the person carrying out the encoding.

The *wit* attribute identifies the witnesses which have the reading in question. It is required if the apparatus gathers together readings from different witnesses, but may be omitted in an apparatus recording the readings of only one witness, e.g. substitutions, divergent opinions on what is in the witness or on how to expand abbreviations, etc. Even in such a one-witness apparatus, however, the *wit* attribute may still be useful when it is desired to record the occurrence of a particular reading in some other witness. For other methods of identifying the witnesses to a reading, see section *13.1.4. Witness Information*.

The *type* attribute allows the encoder to classify readings in any convenient way, for example as substantive variants of the lemma:

```
<app>
  <lem  wit="#El #Hg">Experience</lem>
  <rdg  wit="#La"  type="substantive">Experiment</rdg>
  <rdg  wit="#Ra2"  type="substantive">Eryment</rdg>
</app>
```

or as orthographic variants:

```
<app>
  <lem  wit="#El #Ra2">though</lem>
  <rdg  wit="#La"  type="orthographic">thouh</rdg>
</app>
```

The *varSeq* and *cause* attributes may be used to convey information on the sequence and cause of variation. In the following apparatus fragment, the reading *Eryment* is tagged as sequential to (derived from) the reading *Experiment*, and the cause is given as loss of the abbreviation for *per*.

```
<app>
  <rdg  wit="#La"  varSeq="1">Experiment</rdg>
  <rdg  wit="#Ra2"  cause="abbreviation_loss"
    varSeq="2">Eryment</rdg>
</app>
```

If a manuscript is written in several hands, and it is desired to report which hand wrote a particular reading, the *hand* attribute should be used. For example, in the Munich manuscript containing the *Carmina Burana*, the word *alle* has been changed to *allen*:

```
<l>Swaz hi gât umbe</l>
<l>daz sint alle megede,</l>
<l>die wellent ân man</l>
<l>
  <app>
    <rdg  wit="#Mu"  varSeq="1"  hand="#m1">alle</rdg>
    <rdg  wit="#Mu"  cause="nachgetragen"
      varSeq="2"  hand="#m2">allen</rdg>
  </app>
disen sumer gân.
</l>
```

Similarly, if a witness is hard to decipher, it may be desired to indicate responsibility for the claim that a particular reading is supported by a particular witness. In line 2212a of *Beowulf*, for example, the manuscript is read in different ways by different scholars; the editor Klaeber prints one text, using parentheses to indicate his expansion, and records in the apparatus two different accounts of the manuscript reading, by Zupitza and Chambers:[81]

```
<l>se ðe on
<app>
    <rdg  wit="#Kl">hea(um) h(æþ)e</rdg>
    <rdg  wit="#ms"  source="#Z">heaðo hlæwe</rdg>
    <rdg  wit="#ms"  source="#Cha">heaum hope</rdg>
```

[81] For the sake of legibility in the example, long marks over vowels are omitted.

```
  </app>
</l>
<l>hord beweotode,</l>
```

Because the *hand* attribute indicates a particular manuscript hand, it is intelligible only on a reading from a single witness. If an encoder wishes to indicate that a particular reading from a list in *wit* is in a particular hand, the <witDetail> element should be used; see section *13.1.4. Witness Information*.

Where there is a greater weight of editorial discussion and interpretation than can conveniently be expressed through the attributes provided on these elements (for example where the editor wishes to discuss how a section of text might be punctuated) this information can be attached to the apparatus in a note.

The <note> element may also be used to record the specific wording of notes in the apparatus of the source edition, as here in a transcription of Friedrich Klaeber's note on *Beowulf* 2207a:

```
<l  n="2207a">syððan <app>
    <lem>Beowulfe</lem>
    <note  source="#Kl">Fol. 179a <mentioned>beowulfe</mentioned>.
        Folio 179, with the last page (Fol. 198b), is the worst part of the
        entire MS. It has been freshened up by a later hand, but not always
        correctly. Information on doubtful readings is in the notes of
        Zupitza and Chambers.</note>
  </app>
</l>
<l  n="2207b">brade rice</l>
```

Notes providing details of the reading of one particular witness should be encoded using the specialized <witDetail> element described in section *13.1.4. Witness Information*.

Encoders should be aware of the distinct fields of use of the attribute values *wit*, *hand*, and *source*. Broadly, *wit* identifies the physical entity in which the reading is found (manuscript, clay tablet, papyrus, printed edition); *hand* refers to the agent responsible for inscribing that reading in that physical entity (scribe, author, inscriber, hand 1, hand 2); *source* indicates the scholar responsible for asserting the existence of that reading in that physical entity. In some cases, the categories may blur: a scholar may produce an edition introducing readings for which he or she is responsible; that edition may itself become a witness in a later critical apparatus. Thus, readings introduced as corrections in the earlier edition will be seen in the later apparatus as witnessed by the earlier edition. As observed in the discussion concerning the discrimination of *hand* and *resp* in transcription of primary sources in section *12.3.2.2. Hand, Responsibility, and Certainty Attributes*, the division of layers of responsibility through various scholars for particular aspects of a particular reading may require the more complex mechanisms for assigning responsibility described in chapter *22. Certainty, Precision, and Responsibility*.

## 13.1.3 Indicating Subvariation in Apparatus Entries

The <rdgGrp> element may be used to group readings, either because they have identical values on one or more attributes, or because they are seen as forming a self-contained variant sequence, or for some other reason. This grouping of readings is entirely optional: no such grouping of readings is required.

`<rdgGrp>` (reading group) within a textual variation, groups two or more readings perceived to have a genetic relationship or other affinity.

The <rdgGrp> element is a member of class `att.textCritical` and therefore can carry the *type*, *cause*, *varSeq*, *hand*, and *resp* attributes described in the preceding section. When values for any of these attributes are given on a <rdgGrp> element, the values given are inherited by the <rdg> or <lem> elements nested within the reading group, unless overridden by a new specification on the individual reading element.

To indicate that both Hg and La vary only orthographically from the lemma, one might tag both readings <rdg type='orthographic'>, as shown in the preceding section. This fact can be expressed more perspicuously, however, by grouping their readings into a <rdgGrp>, thus:

```
<app>
  <lem wit="#El #Ra2">though</lem>
  <rdgGrp type="orthographic">
    <rdg wit="#La">thogh</rdg>
    <rdg wit="#Hg">thouh</rdg>
  </rdgGrp>
</app>
```

Similarly, <rdgGrp> may be used to organize the substantive variants of an apparatus entry. Editors may need to indicate that each of a group of witnesses may be taken as all supporting a particular reading, even though there may be variation concerning the exact form of that reading in, or the degree of support offered by, those witnesses. For example: one may identify three substantive variants on the first word of Chaucer's *Wife of Bath's Prologue* in the manuscripts: these might be expressed in regularized spelling as *Experience*, *Experiment*, and *Eriment*. In fact, the manuscripts display many different spellings of these words, and a scholar may wish both to show that the manuscripts have all these variant spellings and that these variant spellings actually support only the three regularized spelling forms. One may term these variant spellings as 'subvariants' of the regularized spelling forms.

This subvariation can be expressed within an <app> element by gathering the readings into three groups according to the normalized form of their reading. All the readings within each group may be accounted subvariants of the main reading for the group, which may be indicated by tagging it as a <lem> element or as <rdg type='groupBase'>.

In this example, the different subvariants on *Experience*, *Experiment*, and *Eriment* are held within three <rdgGrp> elements nested within the enclosing <app> element:

```
<app type="substantive">
  <rdgGrp type="subvariants">
    <lem wit="#El #Hg">Experience</lem>
    <rdg wit="#Ha4">Experiens</rdg>
  </rdgGrp>
  <rdgGrp type="subvariants">
    <lem wit="#Cp #Ld1">Experiment</lem>
    <rdg wit="#La">Ex<g ref="#per"/>iment</rdg>
  </rdgGrp>
  <rdgGrp type="subvariants">
    <lem resp="#ed2013">Eriment</lem>
    <rdg wit="#Ra2">Eryment</rdg>
  </rdgGrp>
</app>
```

From this, one may deduce that the regularized reading *Experience* is supported by all three manuscripts El Hg Ha4, although the spelling differs in Ha4, and that the regularized reading *Eriment* is supported by Ra2, even though the form differs in that manuscript. Accordingly, an application which recognizes that these apparatus entries show subvariation may then assign all the witnesses instanced as attesting the sub-variants on that lemma as actually supporting the reading of the lemma itself at a higher level of classification. Thus, Ha4 here supports the reading *Experience* found in El and Hg, even though it is spelt slightly differently in Ha4.

Reading groups may nest recursively, so that variants can be classified to any desired depth. Because apparatus entries may also nest, the <app> element might also be used to group readings in the same way. The example above is substantially identical to the following, which uses <app> instead of <rdgGrp>:

```
<app n="a1" type="substantive">
  <rdg wit="#El #Hg #Ha4">
    <app n="a2" type="orthographic">
      <lem wit="#El #Hg">Experience</lem>
      <rdg wit="#Ha4">Experiens</rdg>
    </app>
```

```
  </rdg>
  <rdg wit="#Cp #Ld1 #La">
    <app n="a3" type="orthographic">
      <lem wit="#Cp #Ld1">Experiment</lem>
      <rdg wit="#La">Ex<g ref="#per"/>iment</rdg>
    </app>
  </rdg>
  <rdg wit="#Ra2">
    <app n="a4" type="orthographic">
      <lem resp="#ed2013">Eriment</lem>
      <rdg wit="#Ra2">Eryment</rdg>
    </app>
  </rdg>
</app>
```

This expresses even more clearly than the previous encoding of this material that at the highest level of classification (apparatus entry A1), this variation has three normalized readings, and that the first of these is supported by manuscripts El, Hg, and Ha4; the second by Cp, Ld1, and La; and the third by Ra2. Some encoders may find the use of nested apparatus entries less intuitive than the use of reading groups, however, so both methods of classifying the readings of a variation are allowed.

Reading groups may also be used to bring together variants which form an apparent developmental sequence, and to make clear that other readings are not part of that sequence, as in the following example, which makes clear that the variant sequence *experiment* to *eriment* says nothing about the relative priority of *experiment* and *experience*:

```
<app type="substantive">
  <rdgGrp type="subvariants">
    <lem wit="#El #Hg">Experience</lem>
    <rdg wit="#Ha4">Experiens</rdg>
  </rdgGrp>
  <rdgGrp type="sequence">
    <rdgGrp varSeq="1" type="subvariants">
      <lem wit="#Cp #Ld1">Experiment</lem>
      <rdg wit="#La">Ex<g ref="#per"/>iment</rdg>
    </rdgGrp>
    <rdgGrp varSeq="2"
      cause="abbreviation_loss">
      <lem resp="#ed2013">Eriment</lem>
      <rdg wit="#Ra2">Eryment</rdg>
    </rdgGrp>
  </rdgGrp>
</app>
```

### 13.1.4 Witness Information

A given reading is associated with the set of witnesses attesting it by listing the witnesses in the *wit* attribute on the <rdg> or <lem> element. Special mechanisms, described in the following sections, are needed to associate annotation on a reading with one specific witness among several (section *13.1.4.1. Witness Detail Information*), to transcribe witness information verbatim from a source edition (section *13.1.4.2. Witness Information in the Source*), and to identify the formal lists of witnesses typically provided in the front matter of critical editions (section *13.1.4.3. The Witness List*).

#### 13.1.4.1 Witness Detail Information

When it is desired to give additional information about the reading of a particular witness or witnesses, such as noting that it appears in the margin or was corrected for the reading, that information may be given in a <witDetail> element. This is a specialized note, which can be linked to both a reading and to one or more of the

witnesses for that reading. The link to the reading may be inferred from <witDetail>'s position or made explicit by the *target* attribute which <witDetail> inherits from the attribute class `att.pointing`; the link to the witness, by the *wit* attribute.

`att.pointing` provides a set of attributes used by all elements which point to other elements by means of one or more URI references.

`@target` specifies the destination of the reference by supplying one or more URI References.

`<witDetail>` (witness detail) gives further information about a particular witness, or witnesses, to a particular reading.

`@wit` (witnesses) indicates the sigil or sigla identifying the witness or witnesses to which the detail refers.

Because it annotates an attribute value, <witDetail> cannot be included in the text at the point of attachment; without a *target* attribute, it refers to the closest preceding <lem> or <rdg>. But if there is any ambiguity or if the <witDetail> refers to multiple readings, *target* must be used to point to the reading(s) being annotated. To indicate that the Ellesmere manuscript has an ornamental capital in the word *Experience*, for example, one might write:

```
<app type="substantive">
  <lem wit="#El #Hg">Experience</lem>
  <witDetail wit="#El">Ornamental capital.</witDetail>
  <rdg wit="#Ha4">Experiens</rdg>
</app>
```

This encoding makes clear that the ornamental capital mentioned is in the Ellesmere manuscript, and not in Hengwrt or Ha4.

Like <note>, <witDetail> may be used to record the specific wording of information in the source text, even when the information itself is captured in some more formal way elsewhere. The example from the *Carmina Burana* above (section *13.1.2. Readings*), for example, might be extended thus, to record the wording of the note explaining that the variant reading adds *n* to the original in a second hand:

```
<l>Swaz hi gât umbe</l>
<l>daz sint alle megede,</l>
<l>die wellent ân man</l>
<l>
  <app>
    <rdg wit="#Mu" hand="#m1">alle</rdg>
    <rdg wit="#Mu" hand="#m2">allen</rdg>
    <witDetail wit="#Mu">
      <mentioned>n</mentioned> nachgetragen.
    </witDetail>
  </app>
disen sumer gân.
</l>
```

Feature structures containing information about the text in a witness (whether retroversion, regularization, or other) can also be linked to specific <lem> and <rdg> instances. See chapter *19. Feature Structures*.

#### 13.1.4.2 Witness Information in the Source

Although <witDetail> provides a good way to annotate witness references in *wit*, lists of sigla [82] may be complex enough that it is impractical to use the combination of *wit* and <witDetail>. Moreover, in the transcription of printed critical editions, it may be desirable to retain for future reference the exact form in which the source edition records the witnesses to a particular reading; this is particularly important in cases of ambiguity in the information, or uncertainty as to the correct interpretation. The <wit> element may be used to transcribe such lists of witnesses to a particular reading.

[82] The Latin word *siglum* ( *sign*), pl. *sigla* denotes the abbreviation used in a critical apparatus to indicate a particular witness.

**<wit>** (wit) contains a list of one or more sigla of witnesses attesting a given reading, in a textual variation.
The <wit> list may appear following a <rdg>, <rdgGrp>, or <lem> element in any apparatus entry. <wit> may be used in a way functionally equivalent to *wit* if the sigla therein are wrapped in <ref>s with *target* attributes pointing to a predefined witness. For example

```
<app>
  <lem>Nondum</lem>
  <rdg  wit="#G #P"  xml:id="rdg1.1nundum"
    ana="#orthographical">nundum</rdg>
  <witDetail  wit="#G"  target="#rdg1.1nundum">corr. <ref  target="#G1">G<hi  rend="super">1</hi>
    </ref>
  </witDetail>
</app>
```

which indicates that the reading 'nundum' for 'nondum' is to be found in MSS G (although it is corrected to *nondum* in the primary hand) and P, might be written:

```
<app>
  <lem>Nondum</lem>
  <rdg  wit="#G #P"  ana="#orthographical">nundum</rdg>
  <wit>
    <ref  target="#G">G</ref>(corr. <ref  target="#G1">G<hi  rend="super">1</hi>
    </ref>)<ref  target="#P">P</ref>
  </wit>
</app>
```

This is somewhat more verbose, but accomplishes the same goal. Because *wit* is more succinct, and because it makes the automated verification of correct witness references easier, using *wit* (with <witDetail> when needed) is almost always to be preferred.

#### 13.1.4.3 The Witness List

A list of all identified witnesses should normally be supplied in the front matter of the edition, or in the <sourceDesc> element of its header. This may be given either as a simple bibliographic list, using the <listBibl> element described in *3.12. Bibliographic Citations and References*, or as a <listWit> element, which contains a series of <witness> elements. Each <witness> element may contain a brief characterization of the witness, given as one or more prose paragraphs. If more detailed information about a manuscript witness is available, it should be represented using the <msDesc> element provided by the `msdescription` module; an <msDesc> may appear within a <listBibl>.

Whether information about a particular witness is supplied by means of a <bibl>, <msDesc>, or <witness> element, a unique siglum for this source should always be supplied, using the global *xml:id* attribute. This identifier can then be used elsewhere to refer to this particular witness.

**<listWit>** (witness list) lists definitions for all the witnesses referred to by a critical apparatus, optionally grouped hierarchically.

**<witness>** (witness) contains either a description of a single witness referred to within the critical apparatus, or a list of witnesses which is to be referred to by a single sigil.

**<msDesc>** (manuscript description) contains a description of a single identifiable manuscript or other text-bearing object such as an early printed book.

**<bibl>** (bibliographic citation) contains a loosely-structured bibliographic citation of which the sub-components may or may not be explicitly tagged.

**<listBibl>** (citation list) contains a list of bibliographic citations of any kind.

The minimal information provided by a witness list is thus the set of sigla for all the witnesses named in the apparatus. For example, the witnesses referenced by the examples of this chapter might simply be listed thus:

```
<listWit>
  <witness  xml:id="Chi3"/>
  <witness  xml:id="Ha4"/>
  <witness  xml:id="Ju"/>
  <witness  xml:id="K"/>
  <witness  xml:id="Kb"/>
  <witness  xml:id="Kl"/>
  <witness  xml:id="Kv"/>
  <witness  xml:id="Ld"/>
  <witness  xml:id="Ld1"/>
  <witness  xml:id="Ln"/>
  <witness  xml:id="Mu"/>
  <witness  xml:id="Ry2"/>
  <witness  xml:id="Wa"/>
  <witness  xml:id="X"/>
</listWit>
```

It is more helpful, however, for witness lists to be somewhat more informative: each <witness> element should contain at least a brief prose description of the witness, perhaps including a bibliographic citation, as in the following examples:

```
<listWit>
  <witness  xml:id="El">Ellesmere, Huntingdon Library 26.C.9</witness>
  <witness  xml:id="Hg">Hengwrt, National Library of Wales,
     Aberystwyth, Peniarth 392D</witness>
  <witness  xml:id="Ra2">Bodleian Library Rawlinson Poetic 149
     (see further <ptr  target="http://example.com/msDescs#MSRP149"/>)</witness>
</listWit>
```

As the last example shows, the witness description here may be complemented by a reference to a full description of the manuscript supplied elsewhere, typically as the content of an <msDesc> or <bibl> element. Alternatively, it may contain a whole paragraph of commentary for each witness:

```
<listWit>
  <witness  xml:id="A">die sog. <soCalled>Kleine (oder alte)
        Heidelberger Liederhandschrift</soCalled>.
  <bibl>Universitätsbibliothek Heidelberg col. pal.
        germ. 357. Pergament, 45 Fll. 18,5 × 13,5 cm.</bibl>
     Wahrscheinlich die älteste der drei großen Hss. Sie
  <quote>datiert aus dem 13. Jahrhundert, etwa um 1275. Ihre Sprache
        weist ins Elsaß, evtl. nach Straßburg. Man geht wohl nicht
        fehl, in ihr eine Sammlung aus dem Stadtpatriziat zu sehen</quote>
     (<bibl>
      <author>Blank</author>, [vgl. <ref>Lit. z. Hss. Bd. 2,
            S. 39</ref>] S. 14</bibl>). Sie enthält 34 namentlich
     genannte Dichter. <quote>Zu den Vorzügen von A gehört, daß
        sie kaum je bewußt geändert hat, so daß sie für
        manche Dichter ... oft den besten Text liefert</quote> (so wohl mit
     Recht <bibl>
      <author>v. Kraus</author>
    </bibl>).</witness>
  <witness  xml:id="a">Bezeichnung <bibl>
      <author>Lachmann</author>
    </bibl>s für die von einer 2. Hand auf bl. 40–43
```

```
    geschriebenen Strophen der Hs. A.</witness>
  <witness  xml:id="B">die <soCalled>Weingartner (Stuttgarter)
       Liederhandschrift</soCalled>. <bibl>Württembergische
       Landesbibliothek Stuttgart, HB XIII poetae germanici 1.
       Pergament, 156 Bll. 15 × 11,5 cm; 25 teils ganzseitig,
       teils halbseitige Miniaturen.</bibl> Kaum vor 1306 in Konstanz
    geschrieben. Sie enthält Lieder von 25 namentlich genannten
    Dichtern. (Dazu kommen Gedichte von einigen ungenannten
    bzw. unbekannten Dichtern, ein Marienlobpreis und eine
    Minnelehre.)</witness>
</listWit>
```

It would however generally be preferable to represent such detailed information using an appropriately structured <msDesc> element, as discussed in chapter *11. Manuscript Description.* Note also that if the witnesses being recorded are not manuscripts but printed works, it may be preferable to document them using the standard <bibl> or <biblStruct> elements described in *3.12. Bibliographic Citations and References*, as in this example:

```
<listBibl>
  <bibl  xml:id="bcn_1482">T. Kempis, De la imitació de Jesuchrist e del
    menyspreu del món (trad. Miquel Peres); Barcelona, 1482, Pere
    Posa. Editio princeps.</bibl>
  <bibl  xml:id="val_1491">T. Kempis, Del menyspreu del món (trad. Miquel
    Peres); València, 1491.</bibl>
  <bibl  xml:id="bcn_1518">T. Kempis, Libre del menysprey del món e de la
    imitació de nostre senyor Déu Jesucrist, (trad. Miquel Peres);
    Barcelona, 1518, Carles Amorós. </bibl>
</listBibl>
```

In text-critical work it is customary to refer to frequently occurring groups of witnesses by means of a single common siglum. Such sigla may be documented as pseudo-witnesses in their own right by including a nested witness list within the witness list, which uses the siglum for the group as its identifier, and supplies a fuller name for the group in its optional child <head> element, before listing the other witnesses contained by the group. For example, the Constant Group C of manuscripts comprising witnesses Cp, La, and S12, might be represented as follows:

```
<listWit>
  <witness  xml:id="Ellesmere">Ellesmere, Huntingdon Library 26.C.9</witness>
<!-- ... -->
  <listWit  xml:id="Con">
    <head>Constant Group C</head>
    <witness  xml:id="Cp">Corpus Christi Oxford MS 198</witness>
    <witness  xml:id="La">British Library Lansdowne 851</witness>
    <witness  xml:id="Sl2">British Library Sloane MS 1686</witness>
  </listWit>
</listWit>
```

That the reading *Experiment* occurs in all three manuscripts can now be indicated simply as follows:

```
<rdg  wit="#Con">Experiment</rdg>
```

The more elaborate example below shows both multiple levels of nesting and a strategy for mapping the the *xml:id* of the witness to the siglum which will be displayed to the reader of a derived visualisation:

```
<witness xml:id="Σ">Servius (<abbr type="siglum">Σ</abbr>) = ΔΓ
<listWit>
   <witness xml:id="Δ">
     <abbr type="siglum">Δ</abbr>
     <listWit>
       <witness xml:id="J">
         <abbr type="siglum">J</abbr> = Metens. Bibl. mun. 292, s. IX </witness>
       <witness xml:id="L">
         <abbr type="siglum">L</abbr> = Leid. Bibl. der Rijksuniv. B.P.L. 52, s. VIII
/ IX</witness>
     </listWit>
   </witness>
   <witness xml:id="Γ">
     <abbr type="siglum">Γ</abbr>
     <listWit>
       <witness xml:id="θ">
         <abbr type="siglum">θ</abbr>
         <listWit>
           <witness xml:id="A">
             <abbr type="siglum">A</abbr> = Caroliruh. Bad. Landesbibl. Aug. CXVI, s.
IX2 (Reichenau); de codice A derivati:
           <listWit>
              <witness xml:id="S">
                <abbr type="siglum">S</abbr> = Sangall. Stiftsbibl. 861 + 862, s. IX /
X</witness>
              <witness xml:id="Guelf">
                <abbr type="siglum">Guelf.</abbr> = Guelf. HAB 2546 (44.23 Aug. fol.),
s. XV</witness>
            </listWit>
           </witness>
           <witness xml:id="O">
             <abbr type="siglum">O</abbr> = Oxon. Bodl. Laud lat. 117, s.
XI<hi rend="superscript">2</hi>
           </witness>
         </listWit>
       </witness>
       <witness xml:id="τ">
         <abbr type="siglum">τ</abbr>
         <listWit>
           <witness xml:id="Pa">
             <abbr type="siglum">Pa</abbr> = Paris. BnF lat. 7959, s. IX
(Tours)</witness>
           <witness xml:id="Pc">
             <abbr type="siglum">Pc</abbr> = Paris. BnF lat. 7961, s. X / XI</witness>
           <witness xml:id="Q">
             <abbr type="siglum">Q</abbr>Flor. BML Plut. 45.14, s. IX</witness>
           <witness xml:id="Lb">
             <abbr type="siglum">Lb</abbr> = corrector cod. <ref target="#L">L</ref>
(sup.)</witness>
         </listWit>
       </witness>
       <witness xml:id="γ">
         <abbr type="siglum">γ</abbr>
         <listWit>
           <witness xml:id="E">
             <abbr type="siglum">E</abbr> = Escorial. Bibl. S. Lorenzo T.II.17, s.
IX<hi rend="superscript">2</hi> (Ital. septentrion.) </witness>
```

```
          <witness  xml:id="Pb">
            <abbr  type="siglum">Pb</abbr>Paris. BnF lat. 16236, s. X / XI</witness>
          <witness  xml:id="Y">
            <abbr  type="siglum">Y</abbr> = Trident. Bib. com. 3388 (olim Vind. 72), s.
                     IX<hi  rend="superscript">2</hi>
          </witness>
          <witness  xml:id="M">
            <abbr  type="siglum">M</abbr> = Monac. Bay. Staatsbibl. Clm 6394, s. XI
          </witness>
        </listWit>
       </witness>
       <witness  xml:id="σ">
        <abbr  type="siglum">σ</abbr>
        <listWit>
          <witness  xml:id="W">
            <abbr  type="siglum">W</abbr> = Guelf. HAB 2091, s. XIII<hi  rend="superscript">ex.</h
          </witness>
          <witness  xml:id="N">
            <abbr  type="siglum">N</abbr> = Neap. Bibl. naz. lat. 5 (olim Vind. 27), s.
                     X<hi  rend="superscript">1</hi>
          </witness>
          <witness  xml:id="U">
            <abbr  type="siglum">U</abbr> = Berolin. Staatsbibl. lat. quart. 219, s.
XII</witness>
        </listWit>
       </witness>
       <witness  xml:id="o">
        <abbr  type="siglum">º</abbr> classis codicum de Γ defluentium quibus
lectiones faciliores in contextum contaminatione inferre valde placuit (= PaPcγσ in
A. 9.1-10.190, 10.397-12.162; PaPcγU in A. 12.162-320; PaPcγ in A. 12.320-522); cf.
praef. n. 30.</witness>
     </listWit>
    </witness>
  </listWit>
</witness>
```

Here we have a summary of the witnesses, with their sigla, used in an edition, as is generally found in the *conspectus siglorum* in the front matter of a critical edition. Families are indicated with Greek letters and manuscript witnesses with Latin letters. The siglum for display is always contained in the <abbr> with *type* 'siglum' child of each witness, so it is always easy to retrieve the display siglum for a given identifier reference.

Situations commonly arise where there are many more or less fragmentary witnesses, such that there may be quite distinct groups of witnesses for different parts of a text or collection of texts. One may treat this with distinct <listWit> elements for each different part. Alternatively, one may have a single <listWit> element at the beginning of the file or in its header listing all the witnesses, partial and complete, for the text, with the attestation of fragmentary witnesses indicated within the apparatus by use of the <witStart> and <witEnd> elements described in section *13.1.5. Fragmentary Witnesses*.

If a witness list is provided, it may be unnecessary to give, in each apparatus entry, an exhaustive list of the witnesses which agree with the base text. An application program can—in principle—compare the witnesses given for each variant found with those given in the full list of witnesses, subtracting from this list all the witnesses not active at this point (perhaps because of lacuna, or because they contain a variation on a different, overlapping lemma) and thence calculate all the manuscripts agreeing with the base text. In practice, encoders may find it less error-prone to list all witnesses explicitly in each apparatus entry.

### 13.1.5 Fragmentary Witnesses

If a witness is incomplete (whether a single fragment, a series of fragments, or a relatively complete text with one or more lacunae), it is usually desirable to record explicitly where its preserved portions begin and end. The

following empty tags, which may occur within any <lem> or <rdg> element, indicate the beginning or end of a fragmentary witness or of a lacuna within a witness:

`<witStart>` (fragmented witness start) indicates the beginning, or resumption, of the text of a fragmentary witness.

`<witEnd>` (fragmented witness end) indicates the end, or suspension, of the text of a fragmentary witness.

`<lacunaStart>` (lacuna start) indicates the beginning of a lacuna in the text of a mostly complete textual witness.

`<lacunaEnd>` (lacuna end) indicates the end of a lacuna in a mostly complete textual witness.

These elements constitute the class `model.rdgPart`, members of which are permitted within the elements <lem> and <rdg> when the module defined by this chapter is included in a schema.

Suppose a fragment of a manuscript X of the *Wife of Bath's Prologue* has a physical lacuna, and the text of the manuscript begins with *auctorite*. In an apparatus this might appear thus, distinguished from the reading of other manuscripts by the presence of the <lacunaEnd> element:

```
<app>
  <lem  wit="#El #Hg">Auctoritee</lem>
  <rdg  wit="#La #Ra2">auctorite</rdg>
  <rdg  wit="#X">
    <lacunaEnd/>auctorite</rdg>
</app>
```

Alternatively, it may be clearer to record this as

```
<app>
  <lem  wit="#El #Hg">Auctoritee</lem>
  <rdg  wit="#La #Ra2 #X">
    <lacunaEnd  wit="#X"/>auctorite</rdg>
</app>
```

since this shows more clearly that the lacuna and the reading of 'auctorite' both appear in witness X. In some cases, the apparatus in the source may commence recording the readings for a particular witness without its being clear whether the previous absence of readings for this witness is due to a lacuna, or to some other reason. The <witStart> element may be used in this circumstance:

```
<app>
  <lem  wit="#El #Hg">Auctoritee</lem>
  <rdg  wit="#La #Ra2">auctorite</rdg>
  <rdg  wit="#X">
    <witStart/>auctorite</rdg>
</app>
```

## 13.2 Linking the Apparatus to the Text

Three different methods may be used to link a critical apparatus to the text:

1. the location-referenced method,
2. the double-end-point-attached method, and
3. the parallel segmentation method.

Both the location-referenced and the double end-point methods may be used with either *in-line* or *external* apparatus, the former dispersed within the base text, the latter held in some separate location, within or outside the document containing the base text. The parallel segmentation method may only be used for in-line apparatus.

Where an *external* apparatus is used, the <listApp> element provides a useful means of grouping together a series of <app> elements of a specific type, or from a particular source:

`<listApp>` (list of apparatus entries) contains a list of apparatus entries.

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

- *`@type`* characterizes the element in some sense, using any convenient classification scheme or typology.
- *`@subtype`* (subtype) provides a sub-categorization of the element, if needed.

<listApp> elements would normally appear in the <back> of a document, but they may also be placed in any other convenient location.

Any document containing <app> elements requires a <variantEncoding> declaration in the <encodingDesc> element of its TEI header, thus:

`<variantEncoding>` (variant encoding) declares the method used to encode text-critical variants.

- *`@method`* indicates which method is used to encode the apparatus of variants.
- *`@location`* indicates whether the apparatus appears within the running text or external to it.

### 13.2.1 The Location-referenced Method

The location-referenced method of encoding apparatus provides a convenient method for encoding printed apparatus; in this method as in most printed editions, the apparatus is linked to the base text by indicating explicitly only the block of text on which there is a variant (noted usually by a canonical reference scheme, or by line number in the edition, such as A 137 or *Page 15 line 1*).

If the location-referenced method is used for an apparatus stored externally to the base text, the TEI header must have the declaration:

```
<variantEncoding method="location-referenced"
  location="external"/>
```

In the <body> of the document, the base text (here El) will appear:

```
<text>
  <body>
    <div n="WBP" type="prologue">
      <head>The Prologe of the Wyves Tale of Bathe</head>
      <l n="1">Experience though noon Auctoritee</l>
      <l>Were in this world ...</l>
    </div>
  </body>
</text>
```

Elsewhere in the document, or in a separate file, the apparatus will appear. On each <app> element, the *loc* attribute should be specified to indicate where the variant occurs in the base text.

```
<app loc="WBP 1">
  <rdg wit="#La">Experiment</rdg>
  <rdg wit="#Ra2">Eryment</rdg>
</app>
```

If the same text is encoded using in-line storage, the apparatus is dispersed through the base text block to which it refers. In this case, the location of the variant can be read from the line in which it occurs.

```
<variantEncoding method="location-referenced"
  location="internal"/>
<!-- ... -->
```

```
<l  n="1">Experience
<app>
   <rdg  wit="#La">Experiment</rdg>
   <rdg  wit="#Ra2">Eryment</rdg>
  </app>
though noon Auctoritee</l>
<l>Were in this world ...</l>
```

Since the location is not required to be exact, the apparatus for a line might also appear at the end of the line:

```
<l  n="1">Experience though noon Auctoritee
<app>
   <rdg  wit="#La"> Experiment</rdg>
   <rdg  wit="#Ra2"> Eryment</rdg>
  </app>
</l>
<l>Were in this world ...</l>
```

When the apparatus is linked to the text by means of location references, as shown here, it is not possible to find automatically the precise portion of text varied by the readings. In order to show explicitly what portion of the base text is replaced by the variant readings, the <lem> element may be used:

```
<l  n="1">Experience though noon Auctoritee
<app>
   <lem  wit="#El">Experience</lem>
   <rdg  wit="#La">Experiment</rdg>
   <rdg  wit="#Ra2">Eryment</rdg>
  </app>
</l>
<l>Were in this world ...</l>
```

Often the lemma will have no attributes, being simply the 'base text reading' and requiring no qualification, but it may optionally carry the normal attributes, as shown here. Some text critics prefer to abbreviate or elide the lemma, in order to save space or trouble; such practice is not forbidden by these Guidelines, but no recommendations are made for conventions of abbreviating the lemma, whether abbreviation of each word, or suppression of all but the first and last word, etc.

Where it is intended that the apparatus be complete enough to allow the reconstruction of the witnesses (or at least of their non-orthographic variations), simple location-reference methods are unlikely to be as successful as the other two methods, which allow the unambiguous reconstruction of the lemma from the encoding.

### 13.2.2 The Double End-Point Attachment Method

In the double end-point attachment method, the beginning and end of the lemma in the base text are both explicitly indicated. It thus differs from the location-referenced method, in which only the larger span of text containing the lemma is indicated. Double end-point attachment permits unambiguous matching of each variant reading against its lemma. It or the parallel-segmentation method should be used in all cases where this is desired, for example where the apparatus is intended to enable full reconstruction of the text, or of the substantives, of every witness.

When the double end-point attachment method is used, the *from* and *to* attributes of the <app> element are used to indicate the beginning and ending points of the reading in the base text: their values are identifiers which occur at the locations in question. If no other markup is present there, the beginning and ending points should be marked using the <anchor> element defined in chapter *17. Linking, Segmentation, and Alignment.* In cases where it is not possible to insert anchors within the base text (e.g. where the text is on a read-only medium) the beginning and end of the lemma may be indicated by using the 'indirect pointing' mechanisms discussed in chapter *17. Linking, Segmentation, and Alignment.* Explicit anchors are more likely to be reliable, and are therefore to be preferred.

The double end-point attachment method may be used with in-line or external apparatus. In the latter case, the base text (here El) will appear with <anchor> elements inserted at every place where a variant begins or ends (unless some element with an identifier already begins or ends at that point):

```
<variantEncoding method="double-end-point"
  location="external"/>
<!-- ... -->
<div n="WBP" type="prologue">
  <head>The Prologe ... </head>
  <l n="1" xml:id="WBP.1">Experience<anchor xml:id="WBP-A2"/> though noon
Auctoritee</l>
  <l>Were in this world ...</l>
</div>
```

The apparatus will be separately encoded:

```
<app from="#WBP.1" to="#WBP-A2">
  <rdg wit="#La">Experiment</rdg>
  <rdg wit="#Ra2">Eryment</rdg>
</app>
```

No <anchor> element is needed at the beginning of the line, since the *from* attribute can use the identifier for the line as a whole; the lemma is assumed to run from the beginning of the element indicated by the *from* attribute, to the end of that indicated by the *to* attribute. If no value is given for *to*, the lemma runs from the beginning to the end of the element indicated by the *from* attribute.

When the apparatus is encoded in-line, it is dispersed through the base text. Only the beginning of the lemma need be marked with an <anchor>, since the <app> is inserted at the end of the lemma, and itself therefore marks the end of the lemma.

```
<variantEncoding method="double-end-point"
  location="internal"/>
<!-- ... -->
<l n="1" xml:id="wbp.1">Experience
<app from="#wbp.1">
    <rdg wit="#La">Experiment</rdg>
    <rdg wit="#Ra2">Eryment</rdg>
  </app>
though noon Auctoritee</l>
<l>Were in this world ...</l>
```

The lemma need not be repeated within the <app> element in this method, as it may be extracted reliably from the base text. If an exhaustive list of witnesses is available, it will also not be necessary to specify just which manuscripts agree with the base text to enable reconstruction of witnesses. An application will be able to determine the manuscripts that witness the base reading, by noting which witnesses are attested as having a variant reading, and inferring the base text reading for all others after adjusting for fragmentary witnesses and for witnesses carrying overlapping variant readings.

Alternatively, if it is desired to make an explicit record of the attestation of the base text, the <lem> element may be embedded within <app>, carrying the witnesses to the base. Thus

```
<app from="#WBP.1" to="#WBP-A2">
  <lem wit="#El #Hg">Experience</lem>
  <rdg wit="#La">Experiment</rdg>
```

```
  <rdg  wit="#Ra2">Eryment</rdg>
</app>
```

This method is designed to cope with 'overlapping lemmata'. For example, at line 117 of the Wife of Bath's Prologue, the manuscripts Hg (Hengwrt), El (Ellesmere), and Ha4 (British Library Harleian 7334) read:

**Hg** And of so parfit wys a wight ywroght

**El** And for what profit was a wight ywroght

**Ha4** And in what wise was a wight ywroght

In this case, one might wish to record *in what wise was* in Ha4 as a single variant for *of so parfit wys* in Hg, and *was a wight* in El and Ha4 as a variant on *wys a wight* in Hg. This method can readily cope with such difficult situations, typically found in large and complex traditions:

```
<l  xml:id="WBP.117"  n="117"> And
<anchor  xml:id="WBP-A117.1"/> of so parfit
<anchor  xml:id="WBP-A117.2"/> wys
<anchor  xml:id="WBP-A117.3"/> a wight
<anchor  xml:id="WBP-A117.4"/> ywroght
<app  from="#WBP-A117.1"  to="#WBP-A117.3">
   <lem  wit="#Hg">of so parfit wys</lem>
   <rdg  wit="#Ha4">in what wise was</rdg>
  </app>
  <app  from="#WBP-A117.2"  to="#WBP-A117.4">
   <lem  wit="#Hg">wys a wight</lem>
   <rdg  wit="#El #Ha4">was a wight</rdg>
  </app>
</l>
```

The parallel segmentation method, to be discussed next, cannot handle overlaps among variants, and would require the individual variants to be split into pieces.

Because creation and interpretation of double end-point attachment apparatus will be lengthy and difficult it is likely that they will usually be created and examined by scholars only with mechanical assistance.

### 13.2.3 The Parallel Segmentation Method

This method differs from the double end-point attachment method in that all variants at any point of the text are expressed as variants on one another. In this method, no two variations can overlap, although they may nest. The texts compared are divided into matching segments all synchronized with one another. This permits direct comparison of any span of text in any witness with that in any other witness. With a positive apparatus, it is straightforward for an application to extract the full text of any one witness from the apparatus.[83]

This method will (by definition) always be satisfactory when there are just two texts for comparison (assuming they are in the same language and script). It will however be less convenient for textual traditions where establishing a base text with variations from it is not a satisfactory goal for the edition, or in some cases where every detail of variation needs to be modeled.

In the parallel segmentation method, each segment of text on which there is variation is marked by an <app> element. If there is a preferred (or base) reading it is tagged with <lem>; each reading is given in a <rdg> element:

```
<variantEncoding  method="parallel-segmentation"
  location="internal"/>
<!-- ... -->
```

[83] Some care must be taken with this approach, as a derived view of a witness may not be a complete and accurate transcription of that witness. It is more likely to be the base text with all readings from that witness applied.

```
<l  n="1">
  <app>
    <lem  wit="#El #Hg">Experience</lem>
    <rdg  wit="#La">Experiment</rdg>
    <rdg  wit="#Ra2">Eryment</rdg>
  </app> though noon Auctoritee
</l>
<l>Were in this world ...</l>
```

This method cannot be used with external apparatus: it must be used in-line. Note that apparatus encoded with this method may be translated into the double end-point attachment method and back without loss of information. Where double-end-point-attachment encodings have no overlapping lemmata, translation of these to the parallel segmentation encoding and back will also be possible without loss of information.

As noted, apparatus entries may nest in this method: if an imaginary fifth manuscript of the text read *Auctoritee, though none experience*, the variation on the individual words of the line would nest within that for the line as a whole:

```
<l  n="1">
  <app>
    <rdg  wit="#Chi3">Auctoritee, though none experience</rdg>
    <rdg>
      <app>
        <rdg  wit="#El #Hg">Experience</rdg>
        <rdg  wit="#La">Experiment</rdg>
        <rdg  wit="#Ra2">Eryment</rdg>
      </app>
      <app>
        <rdg  wit="#El #Ra2">though</rdg>
        <rdg  wit="#Hg">thogh</rdg>
        <rdg  wit="#La">thouh</rdg>
      </app>
      <app>
        <rdg  wit="#El #Hg">noon Auctorite</rdg>
        <rdg  wit="#La #Ra2">none auctorite</rdg>
      </app>
    </rdg>
  </app>
</l>
```

Parallel segmentation cannot, however, deal very gracefully with variants which overlap without nesting: such variants must be broken up into pieces in order to keep all witnesses synchronized.

### 13.2.4 Other Linking Methods

When an apparatus is provided it does not need to be given at the location in the transcription where the variation, emendation, attribution, or other apparatus observation occurs. Instead it may be stored in a separate place in the same file, or indeed in another file, and point to the location at which it is meant to be used. Storing apparatus entries separately can be beneficial when encoding multiple competing, potentially overlapping, interpretations of the same point in the source texts.

The location-referenced method can be used to point a position in a text using the *loc* attribute and a canonical reference that is understood and documented in the context of the file where it is used. Where possible it is recommended that other methods use the *from* attribute to point to an *xml:id* attribute on an <anchor> or other element at the location where the apparatus observation takes place. The contents of an element pointed to are understood to be equivalent to a <lem> if none exists in the <app>, and if a <lem> does exist this should replace any content.

The *from* attribute is a `teidata.pointer` datatype and thus contains a URI as a value. This means that it can point directly to an *xml:id*, an *xml:id* in another local file, or indeed a file identified by any URL or URN.

```
<l  n="1">
  <seg  xml:id="WBP-so.1.1">Experience</seg> though noon Auctoritee
</l>
<!-- In another file -->
<app  from="example.xml#WBP-so.1.1">
  <rdg  wit="#La">Experiment</rdg>
  <rdg  wit="#Ra2">Eryment</rdg>
</app>
```

This could also be encoded as:

```
<l  n="1">
  <anchor  xml:id="WBP-so.1.1a"/> though noon Auctoritee
</l>
<!-- In another file -->
<app  from="http://www.example.com/example.xml#WBP-so.1.1a">
  <lem>Experience</lem>
  <rdg  wit="#La">Experiment</rdg>
  <rdg  wit="#Ra2">Eryment</rdg>
</app>
```

However, this should be considered more fragile since a full reading of the <lem> is not provided in the source file.

In addition, URLs can contain XPointer schemes including xpath(), range(), and string-range() which can be used in providing the location of an <app> that is stored separately from the text to which it applies. Both *from* and *to* can be used, as in the double end-point attachment method, to identify the starting and ending location for an apparatus using XPointer schemes described in *17.2.4. TEI XPointer Schemes* section to more precisely identify this location where beneficial.

```
<l  n="1"  xml:id="WP.1a">Experience though noon Auctoritee</l>
<!-- In another file -->
<app  from="example.xml#string-range(WP.1a,0,10)">
  <lem>Experience</lem>
  <rdg  wit="#La">Experiment</rdg>
  <rdg  wit="#Ra2">Eryment</rdg>
</app>
```

If only the *from* attribute is provided then it should be understood that this supplies the location of the textual variance that the apparatus documents. If the *from* attribute contains an XPointer scheme that identifies a range of text (or elements) then this is understood to record the starting and ending of the range as in the double end-point attachment method. In such a case a @to attribute is unnecessary.

## 13.3 Using Apparatus Elements in Transcriptions

It is often desirable to record different transcriptions of one stretch of text. These variant transcriptions may be grouped within a single <app> element. An application may then construct different 'views' of the transcription by extraction of the appropriate variant readings from the apparatus elements embedded in the transcription.

For example, alternative expansions can be recorded in several different <expan> elements, all grouped within an <app> element. Consider, for example, the three different transcriptions given below of line 105 of the Hengwrt manuscript of Chaucer's *The Wife of Bath's Prologue*. The last word of the line *Virginite is grete perfection* is written *perfectio* followed by two minims over which a bar has been drawn, which has been read in different ways by

different scholars. The first transcription, by Elizabeth Solopova, represents the two minims with bar above as a special composite character using the <g> element. This transcription notes this as a mark of abbreviation but gives no expansion for it. A second transcriber, F. J. Furnivall, regards the bar as an abbreviation of *u*, and therefore reads the two minims as an *n*. A third transcriber, P. G. Ruggiers, regards the bar as an abbreviation of *n*, reading the minims as *u*. This information may be held within an <app> structure, as follows:

```
Virginite is grete
<app>
  <rdg  source="#ES">perfectio<am>
     <g  ref="#ii"/>
    </am>
  </rdg>
  <rdg  source="#FJF">perfectio<ex>u</ex>n</rdg>
  <rdg  source="#PGR">perfectiou<ex>n</ex>
  </rdg>
</app>
```

This example uses special purpose elements <am> and <ex> used to represent abbreviation marks and editorial expansion respectively; these elements are provided by the `transcr` module documented in chapter *12. Representation of Primary Sources*, which should be consulted for further discussion of methods of representing multiple readings of a source.

Editorial notes may also be attached to <app> structures within transcriptions. Here, editorial preference for Ruggiers' expansion and an explanation of that preference is given:

```
Virginite is grete
<app>
  <rdg  source="#ES">perfecti<am>
     <g  ref="#ii"/>
    </am>
  </rdg>
  <rdg  xml:id="f105"  source="#FJF">perfectio<ex>u</ex>n</rdg>
  <rdg  xml:id="r105"  source="#PGR">perfectiou<ex>n</ex>
  </rdg>
</app>
<!-- ... <note> appearing elsewhere in the document ... -->
<note  target="#r105 #f105">Furnivall's expansion implies that the bar
is an abbreviation for 'u'. There are no certain instances of
this mark as an abbreviation for 'u' in these manuscripts and it is
widely used as an abbreviation for 'n'. Ruggiers' expansion is to
be accepted.</note>
```

In most cases, elements used to indicate features of a primary textual source may be represented within an <app> structure simply by nesting them within its readings, just as the <am> and <ex> elements are nested within the <rdg> elements in the example just given. However, in cases where the tagged feature extends across a span of text which might itself contain variant readings which it is desired to represent by <app> structures, some adaptation of the tagging may be necessary. For example, a span of text may be marked in the transcription of the primary source as a single deletion but it may be desirable to represent just a few words from this source as individual deletions within the context of a critical apparatus drawing together readings from this and several other witnesses. In this case, the tagging of the span of words as one deletion may need to be decomposed into a series of one-word deletions for encoding within the apparatus. If it is important to record the fact that all were deleted by the same act, the markup may use the <join> element or the *next* and *prev* attributes defined by chapter *17. Linking, Segmentation, and Alignment*.

## 13.4 Strategies for Encoding Variation

Textual variation may manifest itself in many ways. Variation most frequently occurs at the phrase level, but is also common at higher structural levels, such as the verse line, paragraph, or chapter. When these structures are involved, some care must be taken in their encoding to ensure that TEI's Abstract Model is not being broken. It would be an error, for example, to have a  in the <lem>, but a <p> in a <rdg> inside the same apparatus entry, because these structures cannot occur at the same level. Similarly, it is an error if the contents of an apparatus entry place a <p> inside another <p> or an <l> inside an <l>.

Phenomena such as omissions and transpositions in witnesses will require some encoding strategies that differ from those in the examples above. An editor wishing to signal an omission in one witness should encode the omission using an empty <rdg>, thus:

```
<app xml:id="d1e372">
  <lem xml:id="d1e373" source="#Heyworth">
    <l n="18">Hypsipyle uacuo constitit in thalamo:</l>
  </lem>
  <rdg xml:id="d1e376" wit="#J"
    cause="homeoarchon"/>
</app>
```

Notice that in this example, the variation occurs at the unit of the verse line. The scribe of MS J has skipped line 18 (probably by mistake) because, like line 19, it begins with the name "Hypsipyle." If a witness contains an interpolation that the editor does not wish to show in the base text, an empty <lem> should be used, in the same fashion.

Transpositions are harder to encode, because they involve variation that occurs in different locations. A single <app> will therefore not be sufficient, and the variants must be linked. For example, in his edition of Propertius 1.16, Housman printed lines 25-6 after line 32, Heyworth prints them in place. We might encode Heyworth's edition, which records Housman's conjecture despite disagreeing with it, as follows:

```
<app xml:id="app-lem-l25-l26"
  exclude="#app-rdg-Housman-l25-26">
  <lem xml:id="d1e462" source="#Heyworth">
    <l n="25" xml:id="l25">desine iam reuocare tuis periuria verbis,</l>
    <l n="26" xml:id="l26">Cynthia, et oblitos parce movere deos;</l>
  </lem>
</app>
```

and then, after line 32:

```
<app xml:id="app-rdg-Housman-l25-26"
  exclude="#app-lem-l25-l26">
  <rdg xml:id="d1e603" source="#Housman">
    <l copyOf="#l25"/>
    <l copyOf="#l26"/>
  </rdg>
  <note target="#d1e603">Housman put these lines after 32.</note>
</app>
```

Note that both <app>s are linked via the *exclude* attribute, because they are mutually exclusive: if one reading is chosen for display in a reading interface, for example, the other must disappear and vice versa. To avoid repetition, the second pair of lines can make use of the *copyOf* attribute. If they were both transposed and somewhat different, then both sets should be written in full.

Apparatus entries may nest when there is variation at both higher and lower structural levels, e.g.:

```
<app xml:id="d1e275">
  <lem xml:id="d1e277" source="#Heyworth">
    <l n="8">
      <app xml:id="d1e280">
        <lem xml:id="d1e281" wit="#N #Λ">ut</lem>
        <rdg xml:id="d1e283" wit="#A">et</rdg>
        <rdg xml:id="d1e285" source="#Nodell">ac</rdg>
        <note target="#d1e287">perhaps</note>
        <rdg xml:id="d1e287" source="#Heyworth"
          cert="low">quam</rdg>
      </app> formosa nouo quae parat ire uiro.</l>
    <l n="9">
      <app xml:id="d1e294">
        <lem xml:id="d1e295">at</lem>
        <rdg xml:id="d1e297" wit="#A">et</rdg>
      </app> non sic, Ithaci digressu <app xml:id="d1e300">
        <lem xml:id="d1e301">mota</lem>
        <rdg xml:id="d1e303" source="#Graevius">immota</rdg>
      </app>, Calypso</l>
    <l n="10">desertis olim fleuerat aequoribus:</l>
    <l n="11">multos illa dies incomptis maesta capillis</l>
  </lem>
  <rdg xml:id="d1e314" wit="#C"
    cause="homoeoteleuton"/>
  <note target="#d1e314">omits lines 8-11 because of homoeoteleuton.</note>
</app>
```

Here, MS C omits lines 8-11, but there are variations the editor wishes to record in the other witnesses which do have these lines. Therefore, an outer <app> gives the lines in the <lem> and the omission in a <rdg>. Further variation is encoded for lines 8 and 9 using nested <app>s.

## 13.5 Module for Critical Apparatus

The module described in this chapter makes available the following components:

**Module textcrit: Critical Apparatus**

- *Elements defined*: app lacunaEnd lacunaStart lem listApp listWit rdg rdgGrp variantEncoding wit witDetail witEnd witStart witness
- *Classes defined*: att.rdgPart att.textCritical att.witnessed model.rdgLike model.rdgPart

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 14 Names, Dates, People, and Places

This chapter describes a module which may be used for the encoding of names and other phrases descriptive of persons, places, or organizations, in a manner more detailed than that possible using the elements already provided for these purposes in the Core module. In section *3.6. Names, Numbers, Dates, Abbreviations, and Addresses* it was noted that the elements provided in the core module allow an encoder to specify that a given text segment is a proper noun, or a *referring string*, and to specify the kind of object named or referred to only by supplying a value for the *type* attribute. The elements provided by the present module allow the encoder to supply a detailed sub-structure for such referring strings, and to distinguish explicitly between names of persons, places, and organizations.

This module also provides elements for the representation of information about the person, place, or organization to which a given name is understood to refer and to represent the name itself, independently of its application. In simple terms, where the core module allows one simply to represent that a given piece of text is a *name*, this module allows one further to represent a *personal name*, to represent the *person* being named, and to represent the *canonical name* being used. A similar range is provided for names of places and organizations. The main intended applications for this module are in biographical, historical, or geographical data systems such as gazetteers and biographical databases, where these are to be integrated with encoded texts.

The chapter begins by discussing attributes common to many of the elements discussed in the remaining parts of the chapter (*14.1. Attribute Classes Defined by This Module*) before discussing specifically the elements provided for the encoding of component parts of personal names (section *14.2.1. Personal Names*), place names (section *14.2.3. Place Names*) and organizational names (section *14.2.2. Organizational Names*). Elements for encoding personal and organizational data are discussed in section *14.3. Biographical and Prosopographical Data*. Elements for the encoding of geographical data are discussed in section *14.3.4. Places*. Finally, elements for encoding onomastic data are discussed in *14.3.7. Names and Nyms*, and the detailed encoding of dates and times is described in section *14.4. Dates*.

## 14.1 Attribute Classes Defined by This Module

Most of the elements made available by this chapter share some important characteristics which are expressed by their membership in specific attribute classes. Members of the class `att.naming` have specialized attributes which support linkage of a naming element with the entity (person, place, organization) being named; members of the class `att.datable` have specialized attributes which support a number of ways of normalizing the date or time of the data encoded by the element concerned.

### 14.1.1 Linking Names and Their Referents

The class `att.naming` is a subclass of the class `att.canonical`, from which it inherits the following attributes:

`att.canonical` provides attributes that can be used to associate a representation such as a name or title with canonical information about the object being named or referenced.

- `@key` provides an externally-defined means of identifying the entity (or entities) being named, using a coded value of some kind.
- `@ref` (reference) provides an explicit means of locating a full definition or identity for the entity being named by means of one or more URIs.

As discussed in *3.6.1. Referring Strings*, these attributes provide two different ways of associating any sort of name with its referent. For cases where all that is required is to provide some minimal information about the person name, for example their occupation or status, the `att.naming` class also provides a simple *role* attribute. It also provides an additional attribute, which allows the name itself to be associated with a base or canonical form:

`att.naming` provides attributes common to elements which refer to named persons, places, organizations etc.

- `@role` may be used to specify further information about the entity referenced by this name in the form of a set of whitespace-separated values, for example the occupation of a person, or the status of a place.
- `@nymRef` (reference to the canonical name) provides a means of locating the canonical form (*nym*) of the names associated with the object named by the element bearing it.

The encoder may use these attributes in combination as appropriate. For example:

```
That silly man <name  role="politician"  type="person">David Paul Brown</name>
has suffered ...
```

The *ref* attribute should be used wherever it is possible to supply a direct link such as a URI to indicate the location of canonical information about the referent.

```
That silly man <name  ref="#DPB1"  type="person">David Paul Brown</name> has
suffered ...
```

This encoding requires that there exist somewhere a <person> element with the identifier DPB1, which will contain canonical information about this particular person, marked up using the elements discussed in *14.3. Biographical and Prosopographical Data* below. The same element might alternatively be provided by some other document, of course, which the same attribute could refer to by means of a URI, as explained in *17.2. Pointing Mechanisms*:

```
That silly man <name  ref="http://www.example.com/personography.xml#DPB1"
  type="person">David Paul Brown</name> has suffered ...
```

More than one URI may be supplied if the name refers to more than one person. For example, assuming the existence of another <person> element for Mrs Brown, with identifier EBB1, a reference to 'the Browns' might be encoded

```
That wretched pair <name  ref="#DPB1 #EBB1"  type="person">the Browns</name> came
to dine ...
```

The *key* attribute is provided for cases where no such direct link is required: for example because resolution of the reference is carried out by some local convention, or because the encoder judges that no such resolution is necessary. As an example of the first case, a project might maintain its own local database system containing canonical information about persons and places, each entry in which is accessed by means of some system-specific identifier constructed in a project-specific way from the value supplied for the *key* attribute.[84] As an example of the second case, consider the use of well-established codifications such as country or airport codes, which it is probably unnecessary for an encoder to expand further:

```
 I never fly from <name  key="LHR"  type="place">Heathrow Airport</name> to
<name  key="FR"  type="place">France</name>
```

However, as explained in *3.6.1. Referring Strings*, interchange is improved by use of tag URIs in *ref* instead of *key*.

The *nymRef* attribute has a more specialized use, where it is the name itself which is of interest rather than the person, place, or organization being named. See section *14.3.7. Names and Nyms* for further discussion.

All members of the att.naming class inherit the following attributes from the att.global.responsibility class:

att.global.responsibility provides attributes indicating the agent responsible for some aspect of the text, the markup or something asserted by the markup, and the degree of certainty associated with it.

- @*resp* (responsible party) indicates the agency responsible for the intervention or interpretation, for example an editor or transcriber.
- @*cert* (certainty) signifies the degree of certainty associated with the intervention or interpretation.

This enables an encoder to record the agency responsible for a given assertion (for example, the name) and the confidence placed in that assertion by the encoder. Examples are given below.

[84] In the module described by chapter *23. Documentation Elements* a similar method is used to link element descriptions to the modules or classes to which they belong, for example.

### 14.1.2 Dating Attributes

Members of the `att.datable` class share the following attributes:

`att.datable` provides attributes for normalization of elements that contain dates, times, or datable events.

- *`@period`* supplies pointers to one or more definitions of named periods of time (typically <category>s, <date>s, or <event>s) within which the datable item is understood to have occurred.

`att.datable.w3c` provides attributes for normalization of elements that contain datable events conforming to the W3C XML Schema Part 2: Datatypes Second Edition.

- *`@when`* supplies the value of the date or time in a standard form, e.g. yyyy-mm-dd.
- *`@notBefore`* specifies the earliest possible date for the event in standard form, e.g. yyyy-mm-dd.
- *`@notAfter`* specifies the latest possible date for the event in standard form, e.g. yyyy-mm-dd.
- *`@from`* indicates the starting point of the period in standard form, e.g. yyyy-mm-dd.
- *`@to`* indicates the ending point of the period in standard form, e.g. yyyy-mm-dd.

The *when* attribute is used to specify a normalized form for any temporal expression, independently of how it is represented in the text, as in the following example:

```
<date when="1807-06-09">June 9th</date> The period is approaching which will terminate
my present
copartnership. On the <date when="1808-01-01">1st Jany.</date> next, it expires by its
own limitation.
```

The *period* attribute provides a convenient way of associating an event or date with a named period. Its value is a pointer which should indicate some other element where the period concerned is more precisely defined. A convenient location for such definitions is the <taxonomy> element in the <classDecl> (classification declaration) in the <encodingDesc> of a TEI header. A <taxonomy> may contain simply a bibliographic reference to an external definition for it. More usefully, it may also contain a series of <category> elements, each with an identifier and a description. The identifier can then be used as the target for a *period* attribute. For example, a taxonomy of named periods might be defined as follows:

```
<taxonomy xml:id="greekperiods">
  <category xml:id="tyranny">
    <catDesc>Before 510 BC</catDesc>
  </category>
  <category xml:id="classical">
    <catDesc>Between 510 and 323 BC</catDesc>
  </category>
  <category xml:id="hellenistic">
    <catDesc>
      <ref target="http://www.wikipedia.com/wiki/Hellenistic">Hellenistic</ref>.
Commonly treated as <date notBefore="-0323" notAfter="-0031">from the death of
Alexander to the Roman conquest.</date>
    </catDesc>
  </category>
  <category xml:id="roman">
    <catDesc>
      <ref target="http://www.wikipedia.com/wiki/Roman_Empire">Roman</ref>
    </catDesc>
  </category>
  <category xml:id="christian">
    <catDesc> The Christian period technically starts at the birth of Jesus, but in
practice is considered to date from
        the conversion of Constantine in <date when="0312">312 AD</date>. </catDesc>
```

```
  </category>
</taxonomy>
```

With these definitions in place, any datable element may be associated with a specific period:

```
<placeName period="#christian">Stauropolis</placeName>
```

The other dating attributes provided by this class support a wide range of methods of specifying temporal information in a normalized form. The *from* and *to* attributes may be used to express the begining and ending of a period of time, for example:

```
<event xml:id="eMBB" from="1955-12-01"
  to="1956-12-20">
  <label>Montgomery Bus Boycott</label>
  <desc>A political and social protest campaign against the policy of racial
segregation on the public transit system of
    the city of <placeName ref="#MONT">Montgomery</placeName>.</desc>
</event>
```

The *notBefore* and *notAfter* attributes may be used to express a range of possibilities for a particular date (or time). For example the following element, extracted from an imaginary prosopographic entry for Anne Calthorpe, indicates that although the exact date of her death is not known, it can be narrowed down to a particular range: from 22 August 1579 to 28 March 1582, inclusive. Ostensibly the encoder has evidence that Anne Calthorpe was alive on the 22nd of August 1579 and evidence that she was already dead on the 28th of March 1582.

```
<death notBefore="1579-08-22"
  notAfter="1582-03-28"/>
```

Since *when* is used for a particular date or time, *from* and *to* for a duration, and *notBefore* and *notAfter* for a date or time within a range, it makes no sense to use *when* in combination with one or more of the others. Thus these Guidelines at present recommend against the use of *when* in combination with any of *from*, *to*, *notBefore*, or *notAfter*.

The *from* or *to* attributes imply that the temporal expression to which they are attached signifies a duration, so the use of either with *notBefore* or *notAfter* means a duration is indicated.

| | *notBefore* | *from* |
|---|---|---|
| ***notAfter*** | range of possibilities, inclusive | duration from *from* to sometime before *notAfter*, inclusive |
| ***to*** | duration from sometime after *notBefore* to *to*, inclusive | duration from *from* to *to*, inclusive |

Some further self-explanatory examples follow:

```
<birth when="1857-03-15">15 March 1857.</birth>
```

```
<birth notBefore="1857-03-01"
  notAfter="1857-04-30">Some time in March or April of 1857.</birth>
```

```
<residence from="1857-03-01"
  to="1857-04-30">Lived in Amsterdam during March and April of 1857.</residence>
```

```
<date from="1857-03-01"
  notAfter="1857-04-30">From the 1st of March to some time later in March or April of
1857.</date>
```

```
<residence notBefore="1857-03-01"
  to="1857-04-30">From the 1st of March or sometime later to the end of April,
1857.</residence>
```

```
<residence from="1856-03" to="1858-04">From sometime in March of 1856 to sometime in
April of 1858.</residence>
```

Normalization of date and time values permits the efficient processing of data (for example, to determine whether one event precedes or follows another). These examples all use the W3C standard format for representation of dates and times. Further examples, and discussion of some alternative approaches to normalization are given in section *14.4.3. More Expressive Normalizations* below.

## 14.2 Names

### 14.2.1 Personal Names

The core <rs> and <name> elements can distinguish names in a text but are insufficiently powerful to mark their internal components or structure. To conduct nominal record linkage or even to create an alphabetically sorted list of personal names, it is important to distinguish between a family name, a forename and an honorary title. Similarly, when confronted with a string such as 'John, by the grace of God, king of England, lord of Ireland, duke of Normandy and Aquitaine, and count of Anjou', the analyst will often wish to distinguish amongst the various constituent elements present, since they provide additional information about the status, occupation, or residence of the person to whom the name belongs. The following elements are provided for these and related purposes:

`<persName>` (personal name) contains a proper noun or proper-noun phrase referring to a person, possibly including one or more of the person's forenames, surnames, honorifics, added names, etc.

`<surname>` (surname) contains a family (inherited) name, as opposed to a given, baptismal, or nick name.

`<forename>` (forename) contains a forename, given or baptismal name.

`<roleName>` (role name) contains a name component which indicates that the referent has a particular role or position in society, such as an official title or rank.

`<addName>` (additional name) contains an additional name component, such as a nickname, epithet, or alias, or any other descriptive phrase used within a personal name.

`<nameLink>` (name link) contains a connecting phrase or link used within a name but not regarded as part of it, such as *van der* or *of*.

`<genName>` (generational name component) contains a name component used to distinguish otherwise similar names on the basis of the relative ages or generations of the persons named.

In addition to the `att.naming` attributes mentioned above, all of the above elements are members of the class `att.personal`, and thus share the following attributes:

`att.personal` (attributes for components of names usually, but not necessarily, personal names) common attributes for those elements which form part of a name usually, but not necessarily, a personal name.

`@full` indicates whether the name component is given in full, as an abbreviation or simply as an initial.

`@sort` (sort) specifies the sort order of the name component in relation to others within the name.

The <persName> element may be used in preference to the general <name> element irrespective of whether or not the components of the personal name are also to be marked. The element <persName> is synonymous with the element <name type="person">, except that its *type* attribute allows for further subcategorization of the personal name itself, for example as a married, birth, pen, pseudo, or religious name. Consequently the following examples are equivalent:

```
That silly man <rs  ref="tag:projectname.org,2012:DPB1"
  type="person">David Paul Brown</rs> has suffered the furniture of his office to be
seized the third time for
rent.
```

```
That silly man <rs  ref="tag:projectname.org,2012:DPB1"
  type="person">
  <name>David Paul Brown</name>
</rs> has suffered ...
```

```
That silly man <name  ref="tag:projectname.org,2012:DPB1"
  type="person">David Paul Brown</name> has suffered ...
```

```
That silly man <persName  ref="tag:projectname.org,2012:DPB1">David Paul
Brown</persName> has suffered ...
```

The <persName> element is more powerful than the <rs> and <name> elements because distinctive name components occurring within it can be marked as such.

Many cultures distinguish between a family or inherited *surname* and additional personal names, often known as *given names*. These should be tagged using the <surname> and <forename> elements respectively and may occur in any order:

```
<persName>
  <surname>Roosevelt</surname>, <forename>Franklin</forename>
  <forename>Delano</forename>
</persName>
<persName>
  <forename>Franklin</forename>
  <forename>Delano</forename>
  <surname>Roosevelt</surname>
</persName>
```

The *type* attribute may be used with both <forename> and <surname> elements to provide further culture- or project-specific detail about the name component, for example:

```
<persName>
  <forename  type="first">Franklin</forename>
  <forename  type="middle">Delano</forename>
  <surname>Roosevelt</surname>
```

```
</persName>
<persName>
  <forename type="given">Margaret</forename>
  <forename type="unused">Hilda</forename>
  <surname type="birth">Roberts</surname>
  <surname type="married">Thatcher</surname>
</persName>
<persName type="religious">Muhammad Ali</persName>
<persName>
  <forename>Norman</forename>
  <surname type="complex">St John Stevas</surname>
</persName>
```

Values for the *type* attribute are not constrained, and may be chosen as appropriate to the encoding needs of the project. They may be used to distinguish different kinds of forename or surname, as well as to indicate the function a name component fills within the whole. In this example, we indicate that a surname is toponymic, and also point to the specific place name from which it is derived:

```
<persName>
  <forename>Johan</forename>
  <surname type="toponymic" ref="#dystvold">Dystvold</surname>
</persName>
<!-- ... -->
<placeName xml:id="dystvold">Dystvold</placeName>
```

The value complex was suggested above for the not uncommon case where the whole of a surname is composed of several other surname elements. These nested surnames may be individually tagged as well, together with appropriate type values:

```
<persName>
  <forename>Kara</forename>
  <surname type="complex">
    <surname type="paternal">Hattersley</surname>- <surname type="maternal">Smith</surname>
  </surname>
</persName>
```

The *full* attribute may be used to indicate whether a name is an abbreviation, initials, or given in full:

```
<persName>
  <forename full="abb">Maggie</forename>
  <surname>Thatcher</surname>
</persName>
```

These elements may be applied as the encoder considers appropriate, including cases where phrases or expressions are used to stand for surnames or forenames, as in the following:

```
<s>
  <persName>
    <forename>Peter</forename>
    <surname>son of Herbert</surname>
  </persName> gives the king 40 m. for having custody of the land and heir of
```

```
<persName>
   <forename>John</forename>
   <surname>son of Hugh</surname>
  </persName>...
</s>
```

Similarly, patronymics may be treated as forenames, thus:

```
... but it remained for <persName>
  <forename>Snorri</forename>
  <forename>Sturluson</forename>
</persName> to combine the two traditions in cyclic form.
```

When a patronymic is used as a surname, however (e.g. by an individual who otherwise would have no surname, but lives in a culture which requires surnames), it may be tagged as such:

```
Even <persName>
  <forename>Finnur</forename>
  <surname>Jonsson</surname>
</persName> acknowledged the artificiality of the procedure...
```

Alternatively, it may be felt more appropriate to mark a patronymic as a distinct kind of name, neither a forename nor a surname, using the <addName> element:

```
<persName>
  <forename>Egill</forename>
  <addName  type="patronym">Skallagrmsson</addName>
</persName>
```

In the following example, the *type* attribute is used to distinguish a patronymic from other forenames:

```
<persName  ref="tag:projectname.org,2012:pn9">
  <forename  sort="2">Sergei</forename>
  <forename  sort="3"  type="patronym">Mikhailovic</forename>
  <surname  sort="1">Uspensky</surname>
</persName>
```

This example also demonstrates the use of the *sort* attribute common to all members of the `model.persNamePart` class; its effect is to state the sequence in which <forename> and <surname> elements should be combined when constructing a sort key for the name.

Some names include generational or dynastic information, such as a number, or phrases such as 'Junior', or 'the Elder'; these qualifications may also be used to distinguish similarly named but unrelated people. In either case, the <genName> element may be used to distinguish such labels from other parts of the name, as in the following examples:

```
<persName  ref="tag:projectname.org,2012:HEMA1">
  <surname>Marques</surname>
  <genName>Junior</genName>, <forename>Henrique</forename>
</persName>
```

```
<persName>
  <forename>Charles</forename>
  <genName>II</genName>
</persName>
```

```
<persName  xml:lang="de">
  <forename>Rudolf</forename>
  <genName>II</genName>
  <surname>von Habsburg</surname>
</persName>
```

```
<persName>
  <surname>Smith</surname>
  <genName>Minor</genName>
</persName>
```

It is also often convenient to distinguish phrases (historically similar to the generational labels mentioned above) used to link parts of a name together, such as 'von', 'of', 'de' etc. It is often a matter of arbitrary choice whether such components are regarded as part of the surname or not; the <nameLink> element is provided as a means of making clear what the correct usage should be in a given case, as in the following examples:

```
<persName  ref="tag:projectname.org,2012:DUDO1">
  <roleName  type="honorific"  full="abb">Mme</roleName>
  <nameLink>de la</nameLink>
  <surname>Rochefoucault</surname>
</persName>
```

```
<persName>
  <forename>Walter</forename>
  <surname>de la Mare</surname>
</persName>
```

Finally, the <addName> and <roleName> elements are used to mark all name components other than those already listed. The distinction between them is that a <roleName> encloses an associated name component such as an aristocratic or official title which exists in some sense independently of its bearer. The distinction is not always a clear one. As elsewhere, the *type* attribute may be used with either element to supply culture- or application-specific distinctions. Some typical values for this attribute for names in the Western European tradition follow:

**nobility** An inherited or life-time title of nobility such as *Lord*, *Viscount*, *Baron*, etc.

**honorific** An academic or other honorific prefixed to a name e.g. *Doctor*, *Professor*, *Mrs.*, etc.

**office** Membership of some elected or appointed organization such as *President*, *Governor*, etc.

**military** Military rank such as *Colonel*.

**epithet** A traditional descriptive phrase or nick-name such as *The Hammer*, *The Great*, etc.

Note, however, that the *role* a person has in a given context (such as *witness*, *defendant*, etc. in a legal document) should not be encoded using the <roleName> element, since this is intended to mark roles which function as part

of a person's name, not the role of the person bearing the name in general. Information about roles, occupations, etc. of a person are encoded within the <person> element discussed below in *14.3. Biographical and Prosopographical Data*.

Here are some further examples of the usage of these elements:

```
<persName ref="tag:projectname.org,2012:PGK1">
  <roleName type="nobility">Princess</roleName>
  <forename>Grace</forename>
</persName>
```

```
<persName ref="tag:projectname.org,2012:GRMO1"
  type="pseudo">
  <addName type="honorific">Grandma</addName>
  <surname>Moses</surname>
</persName>
```

```
<persName ref="tag:projectname.org,2012:SLWICL1">
  <roleName type="office">President</roleName>
  <forename>Bill</forename>
  <surname>Clinton</surname>
</persName>
```

```
<persName ref="tag:projectname.org,2012:MOGA1">
  <roleName type="military">Colonel</roleName>
  <surname>Gaddafi</surname>
</persName>
```

```
<persName ref="tag:projectname.org,2012:FRTG1">
  <forename>Frederick</forename>
  <addName type="epithet">the Great</addName>
</persName>
```

A name may have any combination of the above elements:

```
<persName ref="tag:projectname.org,2012:EGBR1">
  <roleName type="office">Governor</roleName>
  <forename sort="2">Edmund</forename>
  <forename full="init" sort="3">G.</forename>
  <addName type="nick">Jerry</addName>
  <addName type="epithet">Moonbeam</addName>
  <surname sort="1">Brown</surname>
  <genName full="abb">Jr</genName>.
</persName>
```

Although highly flexible, these mechanisms for marking personal name components will not cater for every personal name, nor for every processing need. Where the internal structure of personal names is highly complex or

where name components are particularly ambiguous, feature structures are recommended as the most appropriate mechanism to mark and analyze them, as further discussed in chapter *19. Feature Structures*.

White space is allowed and therefore significant between elements within <name>, <persName>, <orgName>, and <placeName>. Therefore

```
  <persName> <forename>Mary</forename> <forename>Ann</forename> <nameLink>De</nameLink><surname>M
</persName>
```

encodes 'Mary Ann DeMint' and

```
  <persName> <forename>Mary</forename><forename>Ann</forename> <nameLink>De</nameLink>
<surname>Mint</surname> </persName>
```

encodes 'MaryAnn De Mint'. See *1.3.1.1.6. XML Whitespace* for more information on whitespace in XML.

### 14.2.2 Organizational Names

In these Guidelines, we use the term 'organization' for any named collection of people regarded as a single unit. Typical examples include institutions such as 'Harvard College' or 'the BBC' and businesses such as 'Apple' or 'Google' but also racial or ethnic groupings or political factions where these are regarded as forming a single agency such as 'the Scythians' or 'the Militant Tendency'. Giving a loosely-defined group of individuals a name often serves a particular political or social agenda and an analysis of the way such phrases are constructed and used may therefore be of considerable importance to the social historian, even where the objective existence of an 'organization' in this sense is harder to demonstrate than that of (say) a named person. In the case of businesses or other formally constituted institutions, the component parts of an organizational name may help to characterize the organization in terms of its perceived geographical location, ownership, likely number of employees, management structure, etc.

Like names of persons or places, organizational names can be marked up as referring strings or as proper names with the <rs> or <name> elements respectively. The element <orgName> is provided for use where it is desired to distinguish organizational names more explicitly.

`<orgName>` (organization name) contains an organizational name.

This element is a member of the same attribute classes as <persName>, as discussed above in *14.1.1. Linking Names and Their Referents*.

The <orgName> element may be used to mark up any form of organizational name:

```
About a year back, a question of considerable interest
was agitated in the <orgName  type="voluntary"
  ref="tag:projectname.org,2012:PAS1">Pennsyla. Abolition
Society</orgName>
```

This encoding is equivalent to, but more specific than, either of the following representations:

```
About a year back, a question of considerable interest
was agitated in the <rs  ref="tag:projectname.org,2012:PAS1"
  type="org">
  <name>Pennsyla. Abolition Society</name>
</rs>.
```

```
About a year back, a question of considerable interest was agitated in the
<name  ref="tag:projectname.org,2012:PAS1"
  type="org">Pennsyla. Abolition Society</name>.
```

As shown above, like the <rs> and <name> elements, the <orgName> element has a *key* attribute with which an external identifier such as a database key can be assigned to the organization name, and also a *ref* attribute which can be used to point directly to an <org> element containing information about the organization itself (see further *14.3.3. Organizational Data*). Its *type* attribute should be used to characterize the name (rather than the organization), for example as an acronym:

```
Mr Frost will be able to earn an
extra fee from <orgName  type="acronym">BSkyB</orgName> rather than the
<orgName  type="acronym">BBC</orgName>
```

as a phrase:

```
 The feeling in
<country>Canada</country> is one of strong aversion to the <orgName  type="phrase">United
States Government</orgName>,
and of predilection for self-government under the
<orgName  type="phrase">English Crown</orgName>
```

```
<orgName>The Justified Ancients of Mu Mu</orgName>
```

or as a composite of other kinds of name:

```
<orgName  type="partnerNames">
  <surname>Ernst</surname> & <surname>Young</surname>
</orgName>
```

The components of an organization's name may include place names as well as personal names:

```
A spokesman from <orgName  type="regional">
  <orgName>IBM</orgName>
  <country>UK</country>
</orgName> said ...
```

or role names:

```
THE TICKET which you will
receive herewith has been formed by the <orgName>Democratic Whig <name  type="role">party</name>
</orgName> after the most careful deliberation, with a reference to all the great
objects of NATIONAL, STATE, COUNTY and
CITY concern, and with a single eye to the <hi>Welfare and Best Interests of the
Community</hi>.
```

As indicated above, organizational names may also be specified hierarchically particularly where the named organization is itself a department or a branch of a larger organizational entity. 'The Department of Modern History, Glasgow University' is an example:

```
<orgName>
  <orgName>Department of Modern History</orgName>
  <orgName>
    <name type="city">Glasgow</name>
    <name type="role">University</name>
  </orgName>
</orgName>
```

### 14.2.3 Place Names

Like other proper nouns or noun phrases used as names, place names can simply be marked up with the <rs> element, or with the <name> element. For cartographers and historical geographers, however, the component parts of a place name provide important information about the relation between the name and some spot in space and time. They also provide important evidence in historical linguistics.

These Guidelines distinguish three ways of referring to places. A place name (represented using the <placeName> element) may consist of one or more names for hierarchically-organized geo-political or administrative units (see section *14.2.3.1. Geo-political Place Names*). A place named simply in terms of geographical features such as mountains or rivers is represented using the <geogName> element (see section *14.2.3.2. Geographic Names*). Finally, an expression consisting of phrases expressing spatial or other kinds of relationship between other kinds of named place may itself be regarded as a way of referring to a place, and hence as a kind of named place (see section *14.2.3.3. Relative Place Names*).

`<placeName>` (place name) contains an absolute or relative place name.

`<geogName>` (geographical name) identifies a name associated with some geographical feature such as Windrush Valley or Mount Sinai.

As members of the `att.naming` class, all of these elements bear the attributes *key*, *ref*, and *nymRef* mentioned above. These attributes are primarily useful as a means of linking a place name with information about a specific place. Recommendations for the encoding of information about a place, as distinct from its name, are provided in *14.3.4. Places* below.

Like the <persName> element discussed in section *14.2.1. Personal Names*, the <placeName> element may be regarded simply as an abbreviation for the elements <name type="place"> or <rs type="place">. The following encodings are thus equivalent:[85]

```
After spending some time in our <rs ref="tag:projectname.org,2012:NY1"
  type="place">modern <name ref="tag:projectname.org,2012:BA1"
    type="place">Babylon</name>
</rs>, <name ref="tag:projectname.org,2012:NY1"
  type="place">New York</name>, I have proceeded to the <rs ref="tag:projectname.org,2012:PH1"
  type="place">City of Brotherly Love</rs>.
```

```
After spending some time in our <placeName ref="tag:projectname.org,2012:NY1">modern
<placeName ref="tag:projectname.org,2012:BA1">Babylon</placeName>
</placeName>, <placeName ref="tag:projectname.org,2012:NY1">New York</placeName>, I
have proceeded
to the <placeName ref="tag:projectname.org,2012:PH1">City of Brotherly
Love</placeName>.
```

[85] Strictly, a suitable value such as figurative should be added to the two place names which are presented periphrastically in the second version of this example. This would preserve the distinction indicated by the choice of <rs> rather than <name> to encode them in the first version of this example.

#### 14.2.3.1 Geo-political Place Names

A place name may contain text with no indication of its internal structure:

```
<placeName>Rochester, NY</placeName>
```

More usually however, a place name of this kind will be further analysed in terms of its constitutive geo-political or administrative units. These may be arranged in ascending sequence according to their size or administrative importance, for example: 'Rochester, New York', or as a single such unit, for example 'Belgium'. These Guidelines provide a hierarchy of generic element names, each of which may be more exactly specified by means of a *type* attribute:

`<district>` (district) contains the name of any kind of subdivision of a settlement, such as a parish, ward, or other administrative or geographic unit.

`<settlement>` (settlement) contains the name of a settlement such as a city, town, or village identified as a single geo-political or administrative unit.

`<region>` (region) contains the name of an administrative unit such as a state, province, or county, larger than a settlement, but smaller than a country.

`<country>` (country) contains the name of a geo-political unit, such as a nation, country, colony, or commonwealth, larger than or administratively superior to a region and smaller than a bloc.

`<bloc>` (bloc) contains the name of a geo-political unit consisting of two or more nation states or countries.

These elements are all members of the `model.placeNamePart` class, members of which may be used anywhere that text is permitted, including within each other as in the following examples:

```
<placeName>
  <settlement type="city">Rochester</settlement>, <region type="state">New
York</region>
</placeName>
```

```
<placeName ref="tag:projectname.org,2012:LSEA1">
  <country type="nation">Laos</country>, <bloc type="sub-continent">Southeast
Asia</bloc>
</placeName>
```

```
<placeName>
  <district type="arondissement">6ème</district>
  <settlement type="city">Paris, </settlement>
  <country>France</country>
</placeName>
```

#### 14.2.3.2 Geographic Names

Places may also be named in terms of geographic features such as mountains, lakes, or rivers, independently of geo-political units. The <geogName> is provided to mark up such names, as an alternative to the <placeName> element discussed above. For example:

```
<geogName ref="tag:projectname.org,2012:MIRI1"
  type="river">Mississippi River</geogName>
```

In addition to the usual phrase level elements, the <geogName> element may contain the following specialized element:

`<geogFeat>` (geographical feature name) contains a common noun identifying some geographical feature contained within a geographic name, such as valley, mount, etc.

Where the <geogFeat> element is used to characterize the kind of geographic feature being named, the <name> element will generally also be used to mark the associated proper noun or noun phrase:

```
<geogName  ref="tag:projectname.org,2012:MIRI1"
  type="river">
  <name>Mississippi</name>
  <geogFeat>River</geogFeat>
</geogName>
```

A more complex example, showing a variety of practices, follows:

```
The isolated ridge separates two great corridors which run from <name  ref="tag:projectname.org,2
  type="place">Glencoe</name> into <geogName  ref="tag:projectname.org,2012:GLET1"
  type="glen">
  <geogFeat>Glen</geogFeat>
  <name>Etive</name>
</geogName>, the <geogName  ref="tag:projectname.org,2012:LAGA1"
  type="hill">
  <geogFeat  xml:lang="gd">Lairig</geogFeat>
  <name>Gartain</name>
</geogName> and the
<geogName  ref="tag:projectname.org,2012:LAEI1"
  type="hill">
  <geogFeat  xml:lang="gd">Lairig</geogFeat>
  <name>Eilde</name>
</geogName>
```

The Gaelic word *lairig* may be glossed as *sloping hill face*. The most efficient way of including this information in the above encoding would be to create a separate <nym> element for this component of the name and then point to it using the *nymRef* attribute, as further discussed in *14.3.7. Names and Nyms*.

#### 14.2.3.3 Relative Place Names

All the place name specifications so far discussed are *absolute*, in the sense that they define only one place. A place may however be specified in terms of its relationship to another place, for example '10 miles northeast of Paris' or 'near the top of Mount Sinai'. These *relative place names* will contain a place name which acts as a referent (e.g. 'Paris' and 'Mount Sinai'). They will also contain a word or phrase indicating the position of the place being named in relation to the referent (e.g. 'the top of', 'north of'). A distance, possibly only vaguely specified, between the referent place and the place being indicated may also be present (e.g. '10 miles', 'near').

Relative place names may be encoded using the following elements in combination with either a <placeName> or a <geogName> element.

`<offset>` (offset) marks that part of a relative temporal or spatial expression which indicates the direction of the offset between the two place names, dates, or times involved in the expression.

`<measure>` (measure) contains a word or phrase referring to some quantity of an object or commodity, usually comprising a number, a unit, and a commodity name.

Some examples of relative place names are:

```
<placeName  ref="tag:projectname.org,2012:NRPA1">
  <offset>near the top of</offset>
  <geogName>
    <geogFeat>Mount</geogFeat>
    <name>Sinai</name>
```

```
  </geogName>
</placeName>
```

```
<placeName>
  <measure>20 km</measure>
  <offset>north of</offset>
  <settlement type="city">Paris</settlement>
</placeName>
```

If desired, the distance specified may be normalized using the *unit* and *quantity* attributes of <measure>:

```
<placeName ref="tag:projectname.org,2012:Duncan">
  <measure unit="km" quantity="17.7">11 miles</measure>
  <offset>Northwest of</offset>
  <settlement type="city">Providence</settlement>, <region type="state">RI</region>
</placeName>
```

The internal structure of place names is like that of personal names—complex and subject to an enormous amount of variation across time and different cultures. The recommendations in this section should however be adequate for a majority of users and applications; they may be extended using the mechanisms described in chapter *24.3. Customization* to add new elements to the existing classes. When the focus of interest is on the name components themselves, as in place name studies for example, the elements discussed in *14.3.7. Names and Nyms* may also be of use. Alternatively, the meaning structure itself may be represented using feature structures (*19. Feature Structures*).

### 14.2.4 Event Names

The <eventName> element is provided in parallel to other naming mechanisms; it is intended especially for 'canonical' names that are used in texts to mention specific named events, whether real or fictional.

```
<p>On <date when="1719-03-19">Monday</date>, <rs type="person">she</rs> was writing
about the
<eventName ref="#SecondDefPrague">1618 Defenestration of Prague</eventName> which
initiated the
<rs type="event" ref="#ThirtyYearsWar">long war</rs>.</p>
```

`<eventName>` (name of an event) contains a proper noun or noun phrase used to refer to an event.

There are two major applications of <eventName>:

- to indicate names of events mentioned in transcribed text, such as in the above example; and
- to indicate the names of events for which details are being provided in an <event> element.

For the first application, as with most other naming elements, the <eventName> element is a member of the `att.canonical` class, and thus the *ref* (or *key*) attribute may be used to refer to a an <event> element (about which see *14.3.5. Events*) or some other data item that contains further information about the event.

When used as a direct child of <event>, an <eventName> provides one of the names of the event about which details are being provided in its parent <event>. A single event may be known by many names. All, only one, or indeed none of these might be provided in the <event> used to provide information about such an event. If more than one <eventName> is used they may be differentiated by use of *type*, *where*, or dating attributes. In cases where multiple <event> elements describe events that have the same name, the *xml:id* attribute of the <event> or an <idno>, <head>, or <label> sibling of the <eventName> may be used for disambiguation and identification. Typically *xml:id* or <idno> would be used for machine-processable disambiguation, and either <head> or <label> for human readers.

### 14.2.5 Object Names

<objectName> (name of an object) contains a proper noun or noun phrase used to refer to an object.

As with other proper nouns or noun phrases used as names, the names of objects may be marked up simply with the <name> element. For those working with a variety of named objects the <objectName> element provides more flexibility.

```
<p>The <objectName ref="#MinsterLovellJewel">Minster Lovell Jewel</objectName> is
probably the most similar to the
<objectName ref="#AlfredJewel">Alfred Jewel</objectName> and was found in
<placeName ref="#MinsterLovell">Minster
    Lovell</placeName> in <placeName ref="#Oxfordshire">Oxfordshire</placeName> and
is kept at the <orgName ref="#AshmoleanMuseum">Ashmolean Museum</orgName>.</p>
```

The <objectName> element may be used to encode any named object whether or not this is a text-bearing object. The use of <objectName> by itself does not categorize the object referenced, but this may be done further with the *type* and *subtype* attributes or through reference to a <taxonomy>. Additionally, the use of the <objectName> element says nothing about the physical reality of the object – that is whether it is real, fictional, purported, or missing – and this may be one aspect that some may wish to record through the *type* attribute. Where more detailed information is available for a named object the *ref* attribute should be used to point to an <object> element or other source of information about this object. The <objectName> element is intended for named objects; where an object is mentioned through a descriptive phrase but not named explicitly the <rs> element should be used.

## 14.3 Biographical and Prosopographical Data

This module defines a number of special purpose elements which can be used to markup biographical, historical, and prosopographical data. We envisage a number of users and uses for these elements. For example, an encoder may be interested in creating or converting a set of biographical records of the type found in a Dictionary of National Biography. Another use is the creation or conversion of a database-like collection of information about a group of people, such as the people referenced in a marked-up collection of documents, or persons who have served as informants in the creation of spoken corpora. It is also appropriate to use these elements to register information relating to those who have taken part in the creation of a TEI document.

To cater for this diversity, these Guidelines propose a flexible strategy, in which encoders may choose for themselves the approach appropriate to their needs. If one were interested, for example, in converting existing DNB-type records, and wanted to preserve the text as is, the <person> element (see *14.3.2. The Person Element*) could simply contain the text of an article, placed within <p> elements, possibly using elements such as <name> or <date> to mark up features of that text. For a more structured entry, however, one would extract the data and place information contained in the text, and encode it directly using the more specific elements described in this section.

### 14.3.1 Basic Principles

Information about people, places, organizations, and events, of whatever type, essentially comprises a series of statements or assertions relating to:

- characteristics or *traits* which do not, by and large, change over time
- characteristics or *states* which hold true only at a specific time
- *events* or incidents which may lead to a change of state or, less frequently, trait,
- external resources where other information on the subject can be found.

'Characteristics' or 'traits' are typically independent of an individual's volition or action and can be either physical, such as hair and eye colour; or cultural, such as ethnicity and caste. The distinction is not entirely straightforward: gender and sex can be regarded as culturally determined with physical manifestations, interpreted based on the social hierarchies of a time and place. Also the division of humanity into different 'races', formalized in the nineteenth century to organize a hierarchy of humanity on the basis of physical characteristics such as skin colour, hair type, and skull measurements, is now recognized as a colonialist social construct.[86] When such a construct defined

[86]For example, see Renato G. Mazzolini's historical survey of European empires and their unstable constructions of race.?

individuals' identities beyond their volition, the markup and documentation should reflect this as a *culturally-determined* characteristic.

'States' include, for example, marital status, place of residence, and position or occupation. Such states have a definite duration, that is, they have a beginning and an end and are typically a consequence of the individual's own action or that of others.

'Changes in state' refer to life-changing or identity-changing events such as birth, marriage, or appointment to office; such events will normally be associated with a specific date or a fairly narrow date-range. Changes in states can also cause or be caused by changes in characteristics. Any statement or assertion about these aspects of a person's life will be based on some source, possibly multiple sources, possibly contradictory. Taking all this into account it follows that each such statement or assertion needs to be documentable, put into a time frame, and be relatable to other statements or assertions.

The elements defined by the module described in this chapter may, for the most part, all be regarded as specializations of one or other of the above three classes. Generic elements for state, trait, and event are also defined:

`<state>` (state) contains a description of some status or quality attributed to a person, place, or organization often at some specific time or for a specific date range.

`<trait>` (trait) contains a description of some status or quality attributed to a person, place, or organization typically, but not necessarily, independent of the volition or action of the holder and usually not at some specific time or for a specific date range.

`<event>` (event) contains data relating to anything of significance that happens in time.

`att.locatable` provides attributes for referencing locations by pointing to entries in a canonical list of places.

- `@where` indicates one or more locations by pointing to a <place> element or other canonical description.

`<listEvent>` (list of events) contains a list of descriptions, each of which provides information about an identifiable event.

When developing a prosopography record of a named entity it is a common practice to refer explicitly to other resources, for example the Library of Congress Name Authority File, Virtual Internationl Authority File (VIAF), a gazetteer of places like Pleiades, or a printed book.

`<idno>` (identifier) supplies any form of identifier used to identify some object, such as a bibliographic item, a person, a title, an organization, etc. in a standardized way.

- `@type` categorizes the identifier, for example as an ISBN, Social Security number, etc.

Here is a simple example:

```
<place xml:id="Rome">
  <placeName>Rome</placeName>
  <location>
    <geo>41.891775 12.486137</geo>
  </location>
  <idno type="Pleiades">423025</idno>
  <note>capital of the Roman Empire</note>
</place>
```

### 14.3.2 The Person Element

Information about a person, as distinct from references to a person, for example by name, is grouped together within a <person> element. Information about a group of people regarded as a single entity (for example 'the audience' of a performance) may be encoded using the <personGrp> element. Note however that information about a group of people with a distinct identity (for example a named theatrical troupe) should be recorded using the <org> element described in section *14.3.3. Organizational Data* below.

These elements may appear only within a <listPerson> element, which groups such descriptions together, and optionally also describes relationships amongst the people listed.

`<listPerson>` (list of persons) contains a list of descriptions, each of which provides information about an identifiable person or a group of people, for example the participants in a language interaction, or the people referred to in a historical source.

`<listRelation>` provides information about relationships identified amongst people, places, and organizations, either informally as prose or as formally expressed relation links.

One or more <listPerson> elements may be supplied within the <standOff> element (see *17.10. The standOff Container*) or, when used to list the participants in a linguistic interaction, within the <particDesc> (participant description) element in the <profileDesc> element of a TEI header. Like other forms of list, however, <listPerson> can also appear within the body of a text when the module defined by this chapter is included in a schema.

The *type* attribute may be used to distinguish lists of people of different kinds where this is considered convenient:

```
<standOff>
  <listPerson  type="fictional">
    <person  xml:id="person_FAS">
      <persName>Adam Schiff</persName>
      <note>District Attorney for <placeName>Manhattan</placeName> in
            seasons 1 to 10 of <title>Law and Order</title>.</note>
    </person>
    <person  xml:id="person_FML">
      <persName>Mike Logan</persName>
      <note>
        <choice>
          <abbr>NYPD</abbr>
          <expan>New York Police
                    Department</expan>
        </choice> Detective regularly appearing in
            seasons 1 to 5 of <title>Law and Order</title> and seasons 5 to 7
            of <title>Law and Order: Criminal Intent</title>.</note>
    </person>
    <person  xml:id="person_FBS">
      <persName>Benjamin Stone</persName>
      <note>Executive Assistant District Attorney for
      <placeName>Manhattan</placeName> in seasons 1 to 4 of <title>Law
                and Order</title>
      </note>
    </person>
    <person  xml:id="person_FJM">
      <persName>Jack McCoy</persName>
      <note>An Executive Assistant District Attorney then District
            Attorney for <placeName>Manhattan</placeName> in seasons 5 to 10
            of <title>Law and Order</title>, in seasons 1, 9, 11, and 19 of
      <title>Law and Order: Special Victims Unit</title>, and in
            season 1 of <title>Law and Order: Trial by Jury</title>.</note>
    </person>
    <person  xml:id="person_FJR">
      <persName>Jamie Ross</persName>
      <note>An Assistant District Attorney for
      <placeName>Manhattan</placeName> in seasons 7 & 8 of
      <title>Law and Order</title>, and a defense attorney in seasons
            10 & 11, and then a judge in <title>Law and Order: Trial by
               Jury</title>.</note>
    </person>
    <person  xml:id="person_FJF">
      <persName>Joe Fontana</persName>
      <note>
        <choice>
```

```
        <abbr>NYPD</abbr>
        <expan>New York Police
                 Department</expan>
      </choice> Detective regularly appearing
          in seasons 15 & 16 of <title>Law and Order</title>.</note>
    </person>
  </listPerson>
<!-- == == -->
  <listPerson  type="real">
    <person  xml:id="person_RAS">
      <persName>Adam Schiff</persName>
      <note>U.S. Representative from California since 2013.</note>
    </person>
    <person  xml:id="person_RML">
      <persName>Mike Logan</persName>
      <note>Gridiron football player for the Pittsburgh Steelers from
          2001 to 2006.</note>
    </person>
    <person  xml:id="person_RBS">
      <persName>Benjamin Stone</persName>
      <note>Michigan State Senator from 1968 to 1979.</note>
    </person>
    <person  xml:id="person_RJM">
      <persName>Jack McCoy</persName>
      <note>Iowa State Representative from 1955 to 1959.</note>
    </person>
    <person  xml:id="person_RJR">
      <persName>Jamie Ross</persName>
      <note>Broadway actor, with occasional forays into television,
          from 1971 to roughly 2007.</note>
    </person>
    <person  xml:id="person_RJF">
      <persName>Joe Fontana</persName>
      <note>A member of Canada's House of Commons from 1987 to 2006,
          and mayor of London, Ontario from 2010 to 2014.</note>
    </person>
  </listPerson>
</standOff>
```

The <person> element carries several attributes. As a member of the classes `att.global.responsibility`, `att.editLike`, and `att.global.source` class, it carries the usual attributes for providing details about the information recorded for that person, such as its reliability or source:

`att.global.responsibility` provides attributes indicating the agent responsible for some aspect of the text, the markup or something asserted by the markup, and the degree of certainty associated with it.

- @*cert* (certainty) signifies the degree of certainty associated with the intervention or interpretation.
- @*resp* (responsible party) indicates the agency responsible for the intervention or interpretation, for example an editor or transcriber.

`att.editLike` provides attributes describing the nature of an encoded scholarly intervention or interpretation of any kind.

- @*evidence* indicates the nature of the evidence supporting the reliability or accuracy of the intervention or interpretation.

`att.global.source` provides attributes used by elements to point to an external source.

- @*source* specifies the source from which some aspect of this element is drawn.

In addition, a small number of very commonly used personal properties may be recorded using attributes specific to <person> and <personGrp>:

`<person>` (person) provides information about an identifiable individual, for example a participant in a language interaction, or a person referred to in a historical source.

`@role` specifies a primary role or classification for the person.

`@sex` specifies the sex of the person.

`@age` specifies an age group for the person.

`<personGrp>` (personal group) describes a group of individuals treated as a single person for analytic purposes.

These attributes are intended for use where only a small amount of data is to be encoded in a more or less normalized form, possibly for many person elements, for example when encoding basic facts about respondents to a questionnaire. When however a more detailed encoding is required for all kinds of information about a person, for example in a historical gazetteer, then it will be more appropriate to use the elements <age>, <sex> and others described elsewhere in this chapter.

Note that the *age* attribute is not intended to record the person's age expressed in years, months, or other temporal unit. Rather it is intended to record into which age bracket, for the purposes of some analysis, the person falls. A simple (perhaps too simple to be useful) binary classification of age brackets would be child and adult. The actual age brackets useful to various projects are likely to be varied and idiosyncratic, and thus these Guidelines make no particular recommendation as to possible values. Instead, individual projects are recommended to define the values they use in their own customization file, using a declaration like the following:

```
<elementSpec ident="person"
 module="namesdates" mode="change">
 <attList>
  <attDef mode="replace" ident="age">
   <datatype>
    <dataRef key="teidata.enumerated"/>
   </datatype>
   <valList type="closed">
    <valItem ident="child">
     <desc>less than 18 years of age</desc>
    </valItem>
    <valItem ident="adult">
     <desc>18 to 65 years of age</desc>
    </valItem>
    <valItem ident="retired">
     <desc>over 65 years of age</desc>
    </valItem>
   </valList>
  </attDef>
 </attList>
</elementSpec>
```

The above declaration, were it properly placed in a customization file, establishes that the *age* attribute of <person> has only three possible values, child, adult, and retired. For more information on customization see *24.3. Customization*.

The <person> element may contain many sub-elements, each specifying a different property of the person being described. The remainder of this section describes these more specific elements. For convenience, these elements are grouped into three classes, corresponding with the tripartite division outlined above: one for traits, one for states and one for events. Each class may contain specific elements for common types of biographical information, and contains a generic element for other, user-defined, types of information.

All the elements in these three classes belong to the attribute class `att.datable`, which provides the following attributes:

att.datable.w3c provides attributes for normalization of elements that contain datable events conforming to the W3C XML Schema Part 2: Datatypes Second Edition.

- *@when* supplies the value of the date or time in a standard form, e.g. yyyy-mm-dd.
- *@notBefore* specifies the earliest possible date for the event in standard form, e.g. yyyy-mm-dd.
- *@notAfter* specifies the latest possible date for the event in standard form, e.g. yyyy-mm-dd.
- *@from* indicates the starting point of the period in standard form, e.g. yyyy-mm-dd.
- *@to* indicates the ending point of the period in standard form, e.g. yyyy-mm-dd.

as discussed in *14.1. Attribute Classes Defined by This Module* above.

#### 14.3.2.1 Personal Characteristics

The model.persStateLike class contains elements describing physical or socially-constructed characteristics, traits, or states of a person. Members of the class comprise the following specific elements:

<affiliation> (affiliation) contains an informal description of a person's present or past affiliation with some organization, for example an employer or sponsor.

<age> (age) specifies the age of a person.

<education> (education) contains a description of the educational experience of a person.

<faith> (faith) specifies the faith, religion, or belief set of a person.

<floruit> (floruit) contains information about a person's period of activity.

<gender> (gender) specifies the gender identity of a person, persona, or character.

<langKnowledge> (language knowledge) summarizes the state of a person's linguistic knowledge, either as prose or by a list of <langKnown> elements.

<nationality> (nationality) contains an informal description of a person's present or past nationality or citizenship.

<occupation> (occupation) contains an informal description of a person's trade, profession or occupation.

<persName> (personal name) contains a proper noun or proper-noun phrase referring to a person, possibly including one or more of the person's forenames, surnames, honorifics, added names, etc.

<persona> provides information about one of the personalities identified for a given individual, where an individual has multiple personalities.

<persPronouns> (personal pronouns) indicates the personal pronouns used, or assumed to be used, by the individual being described.

<residence> (residence) describes a person's present or past places of residence.

<sex> (sex) specifies the sex of an organism.

<socecStatus> (socio-economic status) contains an informal description of a person's perceived social or economic status.

<state> (state) contains a description of some status or quality attributed to a person, place, or organization often at some specific time or for a specific date range.

<trait> (trait) contains a description of some status or quality attributed to a person, place, or organization typically, but not necessarily, independent of the volition or action of the holder and usually not at some specific time or for a specific date range.

All, apart from <langKnowledge> and <persona>, allow content of ordinary prose containing phrase-level elements.

```
<socecStatus ref="tag:projectname.org,2012:AB1">Status AB1 in the RG Classification
scheme</socecStatus>
```

Encoders should take particular care to be explicit about culturally determined concepts such as nationality, ethnicity, age, or sex/gender system that influence their markup.[87] For example, when recording personal age in different cultures, there may be different assumptions about the point from which age is reckoned. Scholars face

[87] Following Gayle Rubin's influential theory, 'Every society also has a sex/gender system—a set of arrangements by which the biological raw material of human sex and procreation is shaped by human, social intervention and satisfied in a conventional manner, no matter how bizarre some of the conventions may be.?'.

similar issues when marking sex and gender according to cultural systems from times and places not their own. A statement of the practice adopted in a given encoding may usefully be provided in the <editorialDecl> element discussed in *2.3.3. The Editorial Practices Declaration*.

The <langKnowledge> element contains either paragraphs or a number of <langKnown> elements; it may take a *tags* attribute, which provides one or more standard codes or 'tag's for the languages. The <langKnown> element must have a *tag* attribute, which indicates the language with the same kind of 'language tag'. These 'language tags' are discussed in detail in *vi.1 Language Identification*.

Furthermore, the <langKnown> element also has a *level* attribute to indicate the level of the person's competence in the language. It is thus possible either to say:

```
<langKnowledge  tags="ff fr wo en">
  <p>Speaks fluent Fulani, Wolof, and French. Some knowledge of
     English.</p>
</langKnowledge>
```

or

```
<langKnowledge>
  <langKnown  level="fluent"  tag="ff">Fulani</langKnown>
  <langKnown  level="fluent"  tag="wo">Wolof</langKnown>
  <langKnown  level="fluent"  tag="fr">French</langKnown>
  <langKnown  level="basic"  tag="en">English</langKnown>
</langKnowledge>
```

The <persona> element may contain the same component elements as a <person> element. Its function is to document a distinct persona assumed by the <person> element containing it. A person, not necessarily fictional, may take on different personas at different times or in different situations, each persona having different personal characteristics, such as name, age, sex, gender etc. We distinguish a persona, which is a set of characteristics associated with one specific individual, from a role, which is a set of characteristics that many different people can assume. An actor does not change their persona when adopting a different role, but none of the personas associated with one person can properly be associated with another.

The <persPronouns> element may be used to indicate the personal pronouns used, or assumed to be used, by the individual being described. It is common practice in email signatures and biographies, for people to include their preferred personal pronouns along with their name or handle. This allows transgender and gender variant people to express how they wish to be identified, without having to share their gender identity (though some do). Cisgender people have also adopted the practice, which normalizes the idea that a person's personal pronouns should not be inferred by their name, sex, gender, or gender presentation. The <persPronouns> element may be used either in transcribed content to encode a phrase used to indicate preferred personal pronouns, or may be used inside a <person> or <persona> element to indicate the associated pronouns.

For example, the following entry from a hypothetical prosopography lists only the nominative case of the preferred pronouns as identified by Miss Major Griffin-Gracy, a historical figure.

```
<person>
  <persName>
    <forename>Miss Major</forename>
    <surname>Griffin-Gracy</surname>
  </persName>
  <birth  when="1940-10-25"/>
  <gender  value="transWoman"
    evidence="selfIdentification">trans woman</gender>
  <persPronouns  value="she"
    evidence="selfIdentification"/>
```

```
  <note>
    <p>Veteran of the Stonewall Riots. Founder of the
        Griffin-Gracy Educational Retreat and Historical
        Center (the House of GG). Activist and advocate for
        transgender and gender-nonconforming people of
        color.</p>
  </note>
</person>
```

Personal pronouns often occur as part of the closer of an email, post, or other electronic communication.

```
<div  type="email">
  <opener>
    <salute>Dear all,</salute>
  </opener>
  <p>With apologies for length. I'm expanding a schema …</p>
<!-- ... -->
  <closer>
    <lb/>
    <persName>Diane Jakacki, Ph.D.</persName>
    <lb/>
    <roleName>Digital Scholarship Coordinator</roleName>
    <lb/>
    <roleName>Affiliate Faculty in Comparative & Digital Humanities</roleName>
    <lb/>
    <orgName>Bucknell University</orgName>
    <lb/>
    <email>d…@….edu</email>
    <lb/>(<persPronouns>she/her/hers</persPronouns>)
  <lb/>
    <lb/>
    <roleName>Principal Investigator</roleName>,
  <lb/>
    <orgName>LAB Cooperative</orgName> and <orgName>REED London Online</orgName>
    <lb/>
    <roleName>Chair</roleName>, <orgName>ADHO Conference Coordinating
Committee</orgName>
  </closer>
</div>
```

The <sex> and <gender> elements each carry a *value* attribute to give values from a project-internal taxonomy, or an external standard.

```
<sex  value="F">female</sex>
<gender  value="W">woman</gender>
```

As elsewhere, these coded values may be used as an alternative to or normalization of the actual descriptive text contained in the element. The previous example might equally well be given as

```
<sex  value="F"/>
<gender  value="W"/>
```

The generic <trait> and <state> elements are also members of this class,

`<trait>` (trait) contains a description of some status or quality attributed to a person, place, or organization typically, but not necessarily, independent of the volition or action of the holder and usually not at some specific time or for a specific date range.

`<state>` (state) contains a description of some status or quality attributed to a person, place, or organization often at some specific time or for a specific date range.

These elements can be used to extend the range of information supplied about an individual's personal characteristics. Either may contain an optional <label> element, used to provide a human-readable specification for the characteristic concerned and a description of the feature itself supplied within a <desc> element. These may be followed by or one or more <p> elements supplying more detailed information about the trait. In either case, these may be followed by one or more notes or bibliographical references. The *type*, *ref*, and *key* attributes may be used to indicate a fuller definition of the combination of feature and value.

```
<trait type="ethnicity" key="alb">
  <label>Ethnicity</label>
  <desc>Ethnic Albanian.</desc>
</trait>
```

These elements are provided as a simple means of extending the set of descriptive features available in a standardized way. For example, there are no predefined elements for such features as eye or hair colour. If these are to be recorded, they may simply be added as new types of trait:

```
<trait type="physical">
  <label>eye colour</label>
  <desc>blue</desc>
</trait>
<trait type="physical">
  <label>hair colour</label>
  <desc>brown</desc>
</trait>
```

If none of the more specialized elements listed above is appropriate, then a choice must be made between the two generic elements <trait> and <state>. If you wish to distinguish between characteristics that are generally perceived to be transient and those which are generally considered unchanging, use <state> for the former, and <trait> for the latter. It may also be helpful to note that traits are typically, but not necessarily, independent of the volition or action of the holder. If the distinction between state and trait is not considered relevant or useful, use <state>.

The <persName> element is repeatable and can, like all TEI elements, take the attribute *xml:lang* to indicate the language of the content of the element, as well as a *type* attribute to indicate the type of name, whether a nickname, maiden or birth name, alternative form, etc. This is useful in cases where, for example, a person is known by a Latin name and also by any number of vernacular names, many or all of which may have claims to 'authenticity'. In order to ensure uniformity, the method generally employed in the library world has been to accept the form found in some authority file, for example that of the American Library of Congress, as the 'base' or 'neutral' form. Feelings can run high on this matter, however, and people are often reluctant to accept as 'neutral' an overtly foreign form of the name of their local saint or hero. Within the <person> element any number of variant forms of a name can be given, with no prioritization, and hence less likelihood of offence. The Icelandic scholar and manuscript collector Árni Magnússon, to give his name in standard modern Icelandic spelling, is known in Danish as Arne Magnusson, the form which he himself, as a long term resident of Denmark, generally used; there is also a Latinized form, Arnas Magnæus, which he used in his scholarly writings. All three forms can be given, and in any order:

```
<person xml:id="ArnMag">
  <persName xml:lang="is">Árni Magnússon</persName>
  <persName xml:lang="da">Arne Magnusson</persName>
  <persName xml:lang="la">Arnas Magnæus</persName>
</person>
```

At the other extreme, a person may be named periphrastically as in the following example:

```
<person xml:id="simon_son_of_richard2">
  <persName>Simon, son of Richard</persName>
  <residence>
    <placeName>
      <region>Essex</region>
    </placeName>
  </residence>
  <floruit notBefore="1219" notAfter="1223">1219-1223</floruit>
</person>
```

Alternatively, the generic <name> element may be used for all of the naming components in a description. For example, a description of the first living held by the Icelandic clergyman and poet Jón Oddsson Hjaltalín might be tagged as follows:

```
<state type="office" from="1777-04-07"
  to="1780-07-12">
  <p>Jón's first living – which he apparently accepted rather reluctantly – was at
<name type="place">Háls í
       Hamarsfirði</name>, <name type="place">Múlasýsla</name>, to which he was
presented on 7 April 1777. He was
    ordained the following month and spent three years at Háls, but was never happy
there, due largely to the general
    penury in which he was forced to live – a recurrent theme throughout the early
part of his life. In June of 1780
    the bishop recommended that Jón should <q xml:lang="da">promoveres til andet
bedre kald, end det hand hidindtil
       har havt</q>, and on 12 July it was agreed that he should exchange livings with
<name type="person"
     ref="tag:projectname.org,2012:ThorJon">sr. Þórður Jónsson</name> at
<name type="place">Kálfafell á Síðu</name>,
  <name type="place">Skaftafellssýsla</name>.</p>
  <bibl>ÞÍ, Stms I.15, p. 733.</bibl>
  <bibl>ÞÍ, Stms I.17, p.
    102.</bibl>
</state>
```

Similarly, the generic <state> or <trait> element may be used in preference to the more specific elements listed above:

```
<state type="nationality"
  notBefore="2002-01-15">
  <label>Nationality</label>
  <desc>American citizen from 15 January 2002.</desc>
</state>
```

is the same as:

```
<nationality notBefore="2002-01-15">American citizen from 15 January
2002.</nationality>
```

or even:

```
<nationality notBefore="2002-01-15"
  key="US"/>
```

#### 14.3.2.2 Personal Events

This section describes <event>s that relate to one or more <person>s, while the section on Events below describes the use of <event> in more general terms.

Events in a person's history are not characteristics of an individual, but often cause an individual to gain such characteristics, or to enter a new state. Most such events, for example marriage, appointment, promotion, or a journey may be recorded using the generic element <event>, which may be grouped with <listEvent>, and has a content model similar to that of <state> and <trait>. The chief difference is that <event> can include a <placeName> element to identify the name of the place where the event occurred.

Two particular events in a person's life, namely birth and death, are both ubiquitous and usually considered particularly important, and thus may be represented by specialized elements for the purpose:

`<birth>` (birth) contains information about a person's birth, such as its date and place.

`<death>` (death) contains information about a person's death, such as its date and place.

In the following example, we give a brief summary of the wedding of Jane Burden to the English writer, designer, and socialist William Morris, encoded as an <event> element embedded within the <person> element used to record data about Morris, though we could equally well have embedded the <event> element within the <person> element for Burden, or have encoded it independently of either <person> element:

```
<person xml:id="WM">
<!-- ... -->
  <event type="marriage" when="1859-04-26">
    <label>Marriage</label>
    <desc>
      <name type="person" ref="#WM">William Morris</name> and <name type="person"
        ref="http://en.wikipedia.org/wiki/Jane_Burden">Jane Burden</name> were married
at <name type="place">St
          Michael's Church, Ship Street, Oxford</name> on <date when="1859-04-26">26
April 1859</date>. The wedding was
        conducted by Morris's friend <name type="person" ref="#RWD">R. W.
Dixon</name> with <name type="person" ref="#CBF">Charles Faulkner</name> as the best
man. The bride was given away by her father, <name type="person" ref="#RB">Robert
Burden</name>. According to the account that <name type="person"
       ref="http://en.wikipedia.org/wiki/Edward_Burne-Jones">Burne-Jones</name> gave
<name type="person" ref="#JWM">Mackail</name>
     <quote>M. said to Dixon beforehand <said>Mind you don't call her Mary</said> but
he did</quote>. The entry in
        the Register reads: <quote>William Morris, 25, Bachelor Gentleman, 13 George
Street, son of William Morris decd.
          Gentleman. Jane Burden, minor, spinster, 65 Holywell Street, d. of Robert
Burden, Groom.</quote> The witnesses
        were Jane's parents and Faulkner. None of Morris's family attended the
ceremony. Morris presented Jane with a
        plain gold ring bearing the London hallmark for 1858. She gave her husband a
double-handled antique silver
        cup.</desc>
    <bibl>J. W. Mackail, <title>The Life of William Morris</title>, 1899.</bibl>
  </event>
</person>
<person xml:id="RB">
  <persName>Robert Burden</persName>
</person>
<person xml:id="RWD">
```

```
  <persName>R.W. Dixon</persName>
</person>
<person  xml:id="CBF">
  <persName>Charles Faulkner</persName>
</person>
<person  xml:id="EBJ">
  <persName>
    <forename>Edward</forename>
    <surname>Burne-Jones</surname>
  </persName>
</person>
<person  xml:id="JWM">
  <persName>J.W. Mackail</persName>
</person>
```

In this example the *ref* attributes on the various <name> elements point either to an external source or to a <person> element within which other information about the person named may be found. As further discussed below (*14.3.2.3. Personal Relationships*), a <relation> element may then be used to link them in a more meaningful way:

```
<relation  name="spouse"  mutual="#WM #JBM"/>
<relation  name="friend"  mutual="#WM #RWD"/>
<relation  name="parent"  active="#RB"
  passive="#JBM"/>
```

As mentioned above, all these elements, both the specific and the generic, are members of the `att.datable` attribute class, which means they can be limited in terms of time. The following encoding, for example, demonstrates that the person named David Jones changed his name in 1966 to David Bowie:

```
<person  xml:id="DB">
  <persName  notAfter="1966">David Jones</persName>
  <persName  notBefore="1966">David Bowie</persName>
</person>
```

All the generic elements are also members of the `att.global.responsibility` and `att.editLike` classes. These classes make available the attributes *cert*, to indicate the degree of certainty, *resp*, the agency responsible, *evidence*, the nature of the evidence used, and *source*, a pointer to a resource from which the information derives. In this way it is possible, in the case of multiple and conflicting sources, to provide more than one view of what happened, as in the following example:

```
<event  type="birth"  resp="#XYZ"  cert="high">
  <p>Born in <name  type="place">Brixton</name> on 8 January
     1947.</p>
</event>
<event  type="birth"  resp="#ABC"  cert="low">
  <p>Born in <name  type="place">Berkhamsted</name> on 9 January
     1947.</p>
</event>
```

#### 14.3.2.3 Personal Relationships

When the module defined by this chapter is included in a schema, the following two elements may be used to document relationships amongst the persons, places, or organizations identified:

`<listRelation>` provides information about relationships identified amongst people, places, and organizations, either informally as prose or as formally expressed relation links.

`<relation>` (relationship) describes any kind of relationship or linkage amongst a specified group of places, events, persons, objects or other items.

`@name` supplies a name for the kind of relationship of which this is an instance.

`@active` identifies the 'active' participants in a non-mutual relationship, or all the participants in a mutual one.

`@mutual` supplies a list of participants amongst all of whom the relationship holds equally.

`@passive` identifies the 'passive' participants in a non-mutual relationship.

These elements are both members of the `att.typed` class, from which they inherit the *type* and *subtype* attributes in the usual way. The value specified for either attribute on a <listRelation> element is implicitly applicable to all of its child <relation> elements, unless overridden.

A *relationship*, as defined here, may be any kind of describable link between specified participants. A participant (in this sense) might be a person, a place, or an organization. In the case of persons, therefore, a relationship might be a social relationship (such as employer/employee), a personal relationship (such as sibling, spouse, etc.) or something less precise such as 'possessing shared knowledge'. A relationship may be *mutual*, in that all the participants engage in it on an equal footing (for example the 'sibling' relationship); or it may not be if participants are not identical with respect to their role in the relationship (for example, the 'employer' relationship). For non-mutual relationships, only two kinds of role are currently supported; they are named *active* and *passive*. These names are chosen to reflect the fact that non-mutual relations are *directed*, in the sense that they are most readily described by a transitive verb, or a verb phrase of the form *is X of* or *is X to*. The subject of the verb is classed as *active*; the direct object of the verb, or the object of the concluding preposition, as *passive*. Thus parents are 'active' and children 'passive' in the relationship 'parent' (interpreted as *is parent of*); the employer is 'active', the employee 'passive', in the relationship *employs*. These relationships can be inverted: parents are 'passive' and children 'active' in the relationship *is child of*; similarly 'works for' inverts the active and passive roles of 'employs'.

For example:

```
<listRelation>
  <relation  name="parent"  active="#P1 #P2"
    passive="#P3 #P4"/>
  <relation  name="spouse"  mutual="#P1 #P2"/>
  <relation  type="social"  name="employer"
    active="#P1"  passive="#P3 #P4"/>
</listRelation>
```

This example defines the relationships amongst a number of people not further described here; we assume however that each person has been allocated an identifier such as P1, P2, etc. which can be linked to using references such as #P1, #P2, etc. Then the above set of <relation> elements describe the following three relationships amongst the people referenced:

- P1 and P2 are parents of P3 and P4.
- P1 and P2 are linked in a mutual relationship called 'spouse'—that is, P2 is the spouse of P1, and P1 is the spouse of P2.
- P1 has the social relationship 'employer' with respect to P3 and P4.

Relationships within places and organizations are further discussed in the relevant sections below. Relationships between for example organizations and places, or places and persons, may be handled in exactly the same way.

### 14.3.3 Organizational Data

The <org> and <listOrg> elements are used to store data about an organization such as its preferred name, its locations, or key persons within it.

`<org>` (organization) provides information about an identifiable organization such as a business, a tribe, or any other grouping of people.

`<listOrg>` (list of organizations) contains a list of elements, each of which provides information about an identifiable organization.

These elements are intended to be used in a way analogous to the <place> and <person> elements discussed elsewhere in this chapter, that is to provide a unique wrapper element for information about an entity, distinct from references to that entity which are typically encoded using a naming element such as <name type="org"> or <orgName>. The content of a naming element will represent the way an organization is named in a given context; the content of an <org> represents the information known to the encoder about that organization, gathered together in a single place, and independent of its textual realization.

An organization is not the same thing as a list or group of people because it has an identity of its own. That identity may be expressed solely in the existence of a name (for example 'The Scythians'), but is likely to consist in the combination of that name with a number of events, traits, or states which are considered to apply to the organization itself, rather than any of its members. For example, a sports team might be described in terms of its membership (a <listPerson>), its fixtures (a <listPlace>), its geographical affiliation (a <placeName>), or any combination of these. It will also have properties which may be used to categorize it in some way such as the kind of sport played, whether the team is amateur or professional, and so on: these are probably best dealt with by means of the *type* attribute. However, it is the name of the sports team alone which identifies it.

The content model for <org> permits any mixture of generic <state>, <trait>, or <event> elements: the presence of the <orgName> element described in *14.2.2. Organizational Names* is however strongly recommended.

In other respects, the <org> element is used in much the same way as <place> or <person>. An organization may have different names at different times:

```
<org xml:id="fab4">
  <orgName notAfter="1960">The Silver Beetles</orgName>
  <orgName from="1960-08">The Beatles</orgName>
</org>
```

The names of the people making up an organization can also change over time, (if they are known at all). For example:

```
<org xml:id="FAB4">
  <orgName notAfter="1960">The Silver Beetles</orgName>
  <orgName notBefore="1960">The Beatles</orgName>
  <state type="membership" from="1960-08"
    to="1962-05">
    <desc>
      <persName>John Lennon</persName>
      <persName>Paul McCartney</persName>
      <persName>George Harrison</persName>
      <persName>Stuart Sutcliffe</persName>
      <persName>Pete Best</persName>
    </desc>
  </state>
  <state type="membership" notBefore="1963">
    <desc>
      <persName>John Lennon</persName>
      <persName>Paul McCartney</persName>
      <persName>George Harrison</persName>
      <persName>Ringo Starr</persName>
    </desc>
  </state>
</org>
```

An <org> may contain subordinate <org>s:

```
<org  xml:id="OUCS">
  <orgName>Oxford University Computing Services</orgName>
  <org  xml:id="OUCSisg">
    <orgName>Information and Support Group</orgName>
  </org>
  <org  xml:id="OUCSig">
    <orgName>Infrastructure Group</orgName>
    <org  xml:id="OUCSig.nt">
      <orgName>Networking Team</orgName>
    </org>
    <org  xml:id="OUCSig.sdt">
      <orgName>System Development Team</orgName>
    </org>
  </org>
  <org  xml:id="OUCSltg">
    <orgName>Learning Technologies Group</orgName>
  </org>
</org>
```

The following example demonstrates the use of the <listOrg> element to group together a number of <org> elements, each of which is defined solely by means of an informal description, itself containing other names.

```
<p>The TEI institutional hosts are: <listOrg>
   <org  xml:id="bu">
     <orgName>Brown University</orgName>
     <desc>The host contribution is made jointly by the <name  type="project">Brown
University Women Writers
              Project</name> and the <orgName>Brown University Library's Center for
Digital Initiatives</orgName>.</desc>
   </org>
   <org  xml:id="na">
     <orgName>Nancy</orgName>
     <desc>Hosting is provided by a group of institutions located in Nancy, France,
coordinated by
     <orgName>Loria</orgName> and also including <orgName>ATILF</orgName> and
<orgName>INIST</orgName>.</desc>
   </org>
   <org  xml:id="ou">
     <orgName>Oxford University</orgName>
     <desc>Hosting is provided by the <orgName>Research Technologies Service</orgName>
at <orgName>Oxford University
              Computing Services</orgName>.</desc>
   </org>
   <org  xml:id="uv">
     <orgName>University of Virginia</orgName>
     <desc>Virginia's host support comes jointly from the <orgName>Institute for
Advanced Technology in the
              Humanities</orgName> and the <orgName>University of Virginia
Library</orgName>.</desc>
   </org>
  </listOrg>
</p>
```

In a more elaborated version of this example, the organizational names tagged using <orgName> might be linked using the *key* or *ref* attribute to a unique <org> element elsewhere.

### 14.3.4 Places

In *14.2.3. Place Names* we discuss various ways of naming places such as towns, countries, etc. In much the same way as these Guidelines distinguish between the encoding of names for people and the encoding of other data about people, so they also distinguish between the encoding of names for places and the encoding of other data about places. In this section we present elements which may be used to record in a structured way data about places of any kind which might be named or referenced within a text. Such data may be useful as a way of normalizing or standardizing references to particular places, as the raw material for a gazetteer or similar reference document associated with a particular text or set of texts, or in conjunction with any form of geographical information system.

The following elements are provided for this purpose:

`<listPlace>` (list of places) contains a list of places, optionally followed by a list of relationships (other than containment) defined amongst them.

`<place>` (place) contains data about a geographic location.

The `model.placeStateLike` class contains elements describing characteristics of a place which have a definite duration, such as its name. Any member of the `model.placeNamePart` may be used for this purpose, since a <place> element will usually contain at least one, and possibly several, <placeName>-like elements indicating the names associated with it, by different people, in different languages, or at different times.

For example, the modern city of Lyon in France was in Roman times known as Lugdunum. Although the modern and the Roman city are not physically co-extensive, they have significant areas which overlap, and we may therefore wish to regard them as the same place, while supplying both names with an indication of the time period during which each was current.

Places usually have physical locations in addition to names. As with the example of Lyon, the precise geographic location and extent of a place may change over time, and so locations like names may need to be qualified with indications of the time period to which they apply. Locations may be specified in a number of ways: as a set of coordinates defining a point or an area on the surface of the earth, or by providing a description of how the place may be found, usually in terms of other place names. For example, we can identify the location of the Canadian city of London, either by specifying its latitude and longitude, or by specifying that we mean the city called London located in the province called Ontario within the country called Canada.

In addition we may wish to supply a brief characterization of the place identified, for example to state that it is a city, an administrative area such as a country, or a landmark of some kind such as a monument or a battlefield. If our typology of places is simple, the open ended *type* attribute is the easiest way to represent it: so we might say <place type="city">, <place type="battlefield"> etc.

Within the <place> element, the following elements may be used to provide more information about specific aspects of the place in a structured form:

`<placeName>` (place name) contains an absolute or relative place name.

`<location>` (location) defines the location of a place as a set of geographical coordinates, in terms of other named geo-political entities, or as an address.

#### 14.3.4.1 Varieties of Location

A location may be specified in one or more of the following ways:

1. by supplying a string representing its coordinates in some standardized way within a <geo> element, as shown below
2. by supplying one or more place name component elements (e.g. <country>, <settlement> etc.) to place it within a geo-political context
3. by supplying a postal address, e.g. using the <address> element
4. by supplying a brief textual description, e.g. using the <desc> element
5. by using a non-TEI XML vocabulary such as the Geography Markup Language

We give examples of all of these methods in the remainder of this section.

The simplest method of specifying a location is by means of its geographic coordinates, supplied within the <geo> element. This may be used to supply any kind of positional information, using one of the many different geodetic systems available. Such systems vary in their format, in their scope or coverage, and more fundamentally in the reference frame (the 'datum') used for the coordinate system itself. The default recommended by these Guidelines is to supply a string containing two real numbers separated by whitespace, of which the first indicates latitude and

the second longitude according to the 1984 World Geodetic System (WGS84); this is the system currently used by most GPS applications which TEI users are likely to encounter.[88]We might therefore record the information about the place known as 'Lyon' as follows:

```
<place xml:id="LYON1" type="city">
  <placeName notBefore="1400">Lyon</placeName>
  <placeName notAfter="0640">Lugdunum</placeName>
  <location>
    <geo>45.769559 4.834843</geo>
  </location>
</place>
```

Identifying Lyon by its geo-political status as a settlement within a country forming part of a larger political entity, we might represent the same 'place' as follows:

```
<place xml:id="LYON2">
  <placeName notBefore="1400">Lyon</placeName>
  <placeName notAfter="0640">Lugdunum</placeName>
  <location>
    <bloc>EU</bloc>
    <country>France</country>
  </location>
</place>
```

Elements such as <bloc> are specialized forms of <placeName>, as discussed in *14.2.3.1. Geo-political Place Names.*

We may use the same procedure to represent the location of smaller places, such as a street or even an individual building:

```
<place xml:id="BGbldg" type="building">
  <placeName>Brasserie Georges</placeName>
  <location>
    <country key="FR"/>
    <settlement type="city">Lyon</settlement>
    <district type="arrondissement">IIème</district>
    <district type="quartier">Perrache</district>
    <placeName type="street">
      <num>30</num>, Cours de Verdun</placeName>
  </location>
</place>
```

Note the use of the *type* attribute to categorize more precisely both the kind of place concerned (a building) and the kind of name used to locate it, for example by characterizing the generic <district> as an 'arrondissement', or a 'quartier'.

We may also treat imaginary places in the same way:

```
<place xml:id="Atl" type="imaginary">
  <placeName>Atlantis</placeName>
  <location>
    <offset>beyond</offset>
```

[88]See `http://earth-info.nga.mil/GandG/wgs84/index.html`. The most recent revision of this standard is known as the Earth Gravity Model 1996.

```
    <placeName>The Pillars of <persName>Hercules</persName>
    </placeName>
  </location>
</place>
```

A <location> sometimes resembles a set of instructions for finding a place:

```
<place  xml:id="MYF">
  <placeName  notAfter="1969">Yasgur's Farm</placeName>
  <placeName  notBefore="1969">Woodstock Festival Site</placeName>
  <location>
    <measure>one mile</measure>
    <offset>north west of</offset>
    <settlement>Bethel</settlement>
    <region>New York</region>
  </location>
</place>
```

The element <address> may also be used to identify a location in terms of its postal or other address:

```
<place  xml:id="locCA"  type="cemetery">
  <placeName>Protestant Cemetery</placeName>
  <placeName  type="official"  xml:lang="it">Cimitero Acattolico</placeName>
  <location  type="geopolitical">
    <country>Italy</country>
    <settlement>Rome</settlement>
    <district>Testaccio</district>
  </location>
  <location  type="address">
    <address>
      <addrLine>Via Caio Cestio, 6</addrLine>
      <addrLine>00153 Roma</addrLine>
    </address>
  </location>
</place>
```

When, as here, the same place is given multiple locations, the *type* attribute should be used to characterize the kind of location, as a means of indicating that these are alternative ways of identifying the same place, rather than that the place is spread across several locations.

The <location> element may thus identify a place to a greater or lesser degree of precision, using a variety of means: a name, a set of names, or a set of coordinates. The <geo> element introduced earlier is by default understood to supply a value expressed in a specific (and widely used) notation. If a <location> contains more than one <geo>, this is interpreted as being really the same place in the universe, but with different systems used to refer to it. If there is a lack of consensus about the location (of, for example, Camelot), more than one <location> should be used, each with its own <geo>.

By default, the content of <geo> is interpreted as following the standard known as the World Geodetic System (WGS). This may be modified, however, in two ways.

Firstly, the content of the <geo> element can be expressed some other way, that is, according to some different geodetic system. The *decls* attribute is used point to a <geoDecl> element defined in the document header, which describes a different datum.

Secondly, the element <geo> may be redefined to contain markup from a different XML vocabulary which is specifically designed to represent this kind of information. This technique is used throughout these Guidelines where specialized markup is required, for example to embed mathematical expressions or vector graphics, and

is further described and exemplified in *24.3.4. Examples of Modification* . For geographic information, suitable non-TEI vocabularies include:

- the OpenGIS Geography Markup Language (GML) being defined by the OGC[89]
- the Keyhole Markup Language (KML) used by Google Maps[90]

In the following example, we have defined the location of the place 'Lyon' using GML and indicated the two names associated with it at different times:

```
<place  xml:id="locLyon"  type="city"
   xmlns:gml="http://www.opengis.net/gml">
  <placeName  notBefore="1400">Lyon</placeName>
  <placeName  notAfter="0640">Lugdunum</placeName>
  <location>
    <geo>
      <gml:Polygon>
        <gml:exterior>
          <gml:LinearRing> 45.256 -110.45 46.46 -109.48 43.84 -109.86 45.8 -109.2 45.256
-110.45 </gml:LinearRing>
        </gml:exterior>
      </gml:Polygon>
    </geo>
  </location>
</place>
```

A <bibl> element may be used within <location> to indicate the source of the location information.

```
<location>
  <geo>53.226658 -0.541254</geo>
  <bibl>
    <title>Roman Inscriptions of Britain</title>, <idno>262</idno>
  </bibl>
</location>
```

#### 14.3.4.2 Multiple Places

A place may contain other places. This containment relation can be directly modelled in XML: thus we can say that the towns of Vilnius and Kaunas are both in a place called Lithuania (or Lietuva) as follows:

```
<place  xml:id="locLith">
  <country>Lithuania</country>
  <country  xml:lang="lt">Lietuva</country>
  <place>
    <settlement>Vilnius</settlement>
  </place>
  <place>
    <settlement>Kaunas</settlement>
  </place>
</place>
```

This does not, of course, imply that Vilnius and Kaunas are the only places constituting Lithuania; only that they are within it. A separate <place> element may indicate that it is a part of Lithuania by supplying a <relation> element, as discussed below (*14.3.4.4. Relations Between Places*).

[89]The OGC is an international voluntary consensus standards organization whose members maintain the Geography Markup Language standard. The OGC coordinates with the ISO TC 211 standards organization to maintain consistency between OGC and ISO standards work. GML is also an ISO standard (ISO 19136:2007).

[90]See `https://developers.google.com/kml/documentation/`

As a further example, the islands of Mauritius, Réunion, and Rodrigues are collectively known as the Mascarene Islands. Grouped together with Mauritius there are also several smaller offshore islands, with rather picturesque French names. These offshore islands do not however constitute an identifiable place as a whole. One way of representing this is as follows:

```
<place xml:id="locMascarenes"
 type="islandGroup">
 <placeName>Mascarene Islands</placeName>
 <placeName>Mascarenhas Archipelago</placeName>
 <place type="island">
  <placeName>Mauritius</placeName>
  <listPlace type="offshoreIslands">
   <place>
    <placeName>La roche qui pleure</placeName>
   </place>
   <place>
    <placeName>Île aux cerfs</placeName>
   </place>
  </listPlace>
 </place>
 <place type="island">
  <placeName>Rodrigues</placeName>
 </place>
 <place type="island">
  <placeName>Réunion</placeName>
 </place>
</place>
```

Here is a more complex example, showing the variety of names associated at different times and in different languages with a set of hierarchically grouped places—the settlement of Carmarthen Castle, within the town of Carmarthen, within the administrative county of Carmarthenshire, Wales.

```
<place xml:id="wales" type="country">
 <placeName xml:lang="cy">Cymru</placeName>
 <placeName xml:lang="en">Wales</placeName>
 <placeName xml:lang="la">Wallie</placeName>
 <placeName xml:lang="la">Wallia</placeName>
 <placeName xml:lang="fro">Le Waleis</placeName>
 <place xml:id="carmarthenshire"
  type="region">
  <region type="county" xml:lang="en"
   notBefore="1284">Carmarthenshire</region>
  <place xml:id="carmarthen"
   type="settlement">
   <placeName xml:lang="en">Carmarthen</placeName>
   <placeName xml:lang="la"
    notBefore="1090" notAfter="1300">Kaermerdin</placeName>
   <placeName xml:lang="cy">Caerfyrddin</placeName>
   <place xml:id="carmarthen_castle"
    type="castle">
    <settlement>castle of Carmarthen</settlement>
   </place>
  </place>
 </place>
</place>
```

As noted previously, <country>, <region>, and <settlement> are all specializations of the generic <placeName> element; they are not specializations of the <place> element. If it is desired to distinguish amongst kinds of *place* this can only be done by means of the *type* attribute as in the above example.

This use of multiple <place> elements should be distinguished from the (possibly simpler) case where a number of places with some property in common are being grouped together for convenience, for example, in a gazetteer. The <listPlace> element is provided as a means of grouping places together where there is no implication that the grouped elements constitute a distinct place. For example:

```
<place xml:id="pl-c-H" type="county">
  <placeName>Herefordshire</placeName>
  <listPlace type="villages">
    <place xml:id="pl-v-AD">
      <placeName>Abbey Dore</placeName>
      <location>
        <geo>51.969604 -2.893146</geo>
      </location>
    </place>
    <place xml:id="pl-v-AB">
      <placeName>Acton Beauchamp</placeName>
    </place>
<!-- ... -->
  </listPlace>
  <listPlace type="towns">
    <place xml:id="pl-t-H">
      <placeName>Hereford</placeName>
    </place>
    <place xml:id="pl-t-L">
      <placeName>Leominster</placeName>
    </place>
<!-- ... -->
  </listPlace>
</place>
```

#### 14.3.4.3 States, Traits, and Events

There are many different kinds of information which it might be considered useful to record for a place in addition to its name and location, and the categories selected are likely to be very project-specific. As with persons therefore these Guidelines make no claim to comprehensiveness in this context. Instead, the generic <state>, <trait>, and <event> elements defined by this module should be used. Each of these may be customized for particular needs by means of their *type* attribute. These are complemented by a small number of predefined elements of general utility:

`<population>` (population) contains information about the population of a place.

`<climate>` (climate) contains information about the physical climate of a place.

`<terrain>` (terrain) contains information about the physical terrain of a place.

These are all specializations of the generic <trait> element. This element may be used for almost any kind of event in the life of a place; no specialized version of this element is proposed, nor do we attempt to enumerate the possible values which might be appropriate for the *type* attribute on any of these generic elements.

Here is an example, showing how the specific and generic elements may be combined:

```
<place xml:id="IS">
  <placeName xml:lang="en">Iceland</placeName>
  <placeName xml:lang="is">Ísland</placeName>
  <location>
    <geo>65.00 -18.00</geo>
  </location>
```

```
  <terrain>
    <desc>Area: 103,000 sq km</desc>
  </terrain>
  <state type="governance" notBefore="1944">
    <p>Constitutional republic</p>
  </state>
  <state type="governance" notAfter="1944">
    <p>Part of the kingdom of <placeName key="DK">Denmark</placeName>
    </p>
  </state>
  <event type="governance" when="1944-06-17">
    <desc>Iceland became independent on 17 June 1944.</desc>
  </event>
  <state type="governance" from="1944-06-17">
    <p>An independent republic since June 1944</p>
  </state>
</place>
```

In the following example, the <climate> example is used to provided a detailed discussion of this particular aspect of the information available about a particular place:

```
<place xml:id="greece">
  <placeName>Greece</placeName>
  <climate>
    <desc>Greece's climate is divided into three well defined classes:</desc>
    <climate>
      <label>Mediterranean</label>
      <desc>It features mild, wet winters and hot, dry summers. Temperatures rarely
reach extremes, although snowfalls
           do occur occasionally even in <placeName>Athens</placeName>,
<placeName>Cyclades</placeName> or
      <placeName>Crete</placeName> during the winter.</desc>
    </climate>
    <climate>
      <label>Alpine</label>
      <desc>It is found primarily in <placeName>
         <offset>Western</offset> Greece</placeName>
            (<placeName>Epirus</placeName>, <placeName>
         <offset>Central</offset> Greece</placeName>,
      <placeName>Thessaly</placeName>, <placeName>
         <offset>Western</offset> Macedonia</placeName> as well as
            central parts of <placeName>Peloponnesus</placeName> like
<placeName>Achaea</placeName>,
      <placeName>Arcadia</placeName> and parts of <placeName>Laconia</placeName> where
the Alpine range pass
           by)</desc>
    </climate>
    <climate>
      <label>Temperate</label>
      <desc>It is found in <placeName>
         <offset>Central</offset> and <offset>Eastern</offset> Macedonia</placeName> as
            well as in <placeName>Thrace</placeName> at places like
<placeName>Komotini</placeName>,
      <placeName>Xanthi</placeName> and <placeName>
         <offset>northern</offset> Evros</placeName>. It features cold,
            damp winters and hot, dry summers.</desc>
```

```
    </climate>
  </climate>
</place>
```

As the above example shows, <state> and <trait> elements, and others of the same class, can be nested hierarchically within each other. When this is done, values for the *type* attribute are to be understood as cumulatively inherited, as elsewhere in the TEI scheme (for example on <category> or <linkGrp>). In the following example, the outermost <population> element concerns the squirrel population between the dates given. This is then broken down into red and gray squirrel populations, and within that into male and female:

```
<population type="squirrel"
  notBefore="1901" notAfter="1902-01-11" resp="#strabo">
  <population type="red" when="1901-01-10">
    <population type="female">
      <desc>12</desc>
    </population>
    <population type="male">
      <desc>15</desc>
    </population>
  </population>
  <population type="gray" when="1902-01-10"
    cert="high">
    <population type="female">
      <desc>23</desc>
    </population>
    <population type="male" cert="low"
      resp="#biber">
      <desc>45</desc>
    </population>
  </population>
</population>
```

The dating and responsibility attributes here behave slightly differently from the *type* attribute: responsibility is not an additive property, and therefore an element either states it explicitly, or inherits it from its nearest ancestor. Dating is slightly different again, in that a child element may specify a date more precisely than its parent, as in the example above.

An <event> may be related to other named entities, and thus may appear within an <org>, <person>, <personGrp>, <persona>, or <place> element as well as within a <listEvent> or another <event>.

#### 14.3.4.4 Relations Between Places

The <relation> element may also be used to express relationships of various kinds between places, or between places and persons, in much the same way as it is used to express relationships between persons alone. Returning to the Mascarene Islands example cited above, we might define the island group and its constituents separately, but indicate the relationship by means of a <relation> element:

```
<listPlace>
  <place xml:id="MASC">
    <placeName>Mascarene islands</placeName>
    <placeName>Mascarenhas Archipelago</placeName>
  </place>
  <place xml:id="MRU">
    <placeName>Mauritius</placeName>
<!-- ... -->
```

```
  </place>
  <place xml:id="ROD">
    <placeName>Rodrigues</placeName>
  </place>
  <place xml:id="REN">
    <placeName>Réunion</placeName>
  </place>
  <relation name="contains" active="#MASC"
    passive="#ROD #MRU #REN"/>
</listPlace>
```

This 'stand-off' style of representation has the advantage that we can now also represent the fact that a place may be a 'part of' more than one other place; for example, Réunion is part of France, as well as part of the Mascarenes. If we add a declaration for France to the list above:

```
<place type="country" xml:id="FRA">
  <placeName>France</placeName>
</place>
```

we can now model this dual allegiance by means of a <relation> element:

```
<relation name="partOf" active="#REN"
  passive="#FRA #MASC"/>
```

### 14.3.5 Events

`<event>` (event) contains data relating to anything of significance that happens in time.

`<eventName>` (name of an event) contains a proper noun or noun phrase used to refer to an event.

`<listEvent>` (list of events) contains a list of descriptions, each of which provides information about an identifiable event.

An event is a change of state whether real, planned, fictional, or purported about which more information is known. The <eventName> element may be used to encode the proper name of an event, if known. The <head> or <label> element may also appear in <event>, either in lieu of or in addition to the <eventName> element, to provide a human-readable heading or editorially supplied label for the event.

These guidelines do not prescribe what encoders count as events, nor their granularity, nesting, or amount of other information to record in relation to each event. Yet most events can be identified by their time, location, or agents involved (be they organizations, persons, or objects).

An <event> element may be used to record information about a place, person, or organization related to the described event; for this reason the element may appear as content of a <place>, <person>, or <org>. However, it is also possible to describe events independently of either a place, person, or organization. This may be useful not only for encoding texts which describe named events and their <eventName>s, but also in applications as chronologies or lists of significant events (e.g., important battles, legislative milestones, or the stages of construction of a large building, etc.).

The <listEvent> element is a member of the `model.listLike` class, and may therefore appear inside <standOff>, or wherever else lists are permitted, in the same way as the <listPerson>, <listPlace>, <listOrg> etc. elements described elsewhere in this chapter.

```
<standOff>
  <listEvent>
    <event when="1713"
      ref="https://www.canadiana.ca/view/oocihm.9_01832/2">
```

```
     <label>Treaty of Utrecht</label>
<!-- The following example presupposes that the editors of the TEI document decided
to understand legislation or a treaty as an <gi>event</gi>, which itself may be
considered a shortcut for several events that led up to the existence or validity
of the law or treaty (e.g. the signing of a treaty could also be modeled as an event,
and the result on paper could be encoded as an <gi>object</gi> or <gi>bibl</gi>). -->
     <desc>France ceded to Great Britain its claims to the <orgName>Hudson's Bay
Company</orgName> territories in
     <placeName>Rupert's Land</placeName>, <placeName>Newfoundland</placeName>, and
<placeName>Acadia</placeName> and
           recognized British suzerainty over <orgName  type="tribe">the
Iroquois</orgName> but retained its other pre-war
           North American possessions, including
     <placeName  key="PEI">Île-Saint-Jean</placeName>
           (now <placeName  key="PEI">Prince Edward Island</placeName>)...</desc>
   </event>
   <event  when="1774"  key="14-GeoIII-c83">
     <label>Quebec Act</label>
     <desc>This act of the British Parliament guaranteed free practice of the Catholic
faith and restored use of the
           French Civil Code for private matters throughout the Province of Quebec,
which had been expanded in territory
           following the <ref>Treaty of Paris</ref>.</desc>
   </event>
   <event  when="1778"
     ref="http://avalon.law.yale.edu/18th_century/del1778.asp">
     <label>Treaty of Fort Pitt</label>
     <desc>Also known as the <name  type="event">Treaty with the Delawares</name>, this
was the first written treaty
           between the newly formed <orgName>United States</orgName> and any Native
American people, in this case, the
     <orgName  type="tribe">Lenape</orgName> or Delawares.</desc>
   </event>
  </listEvent>
</standOff>
```

The next example demonstrates the use of the *where* attribute with <event>, coordinated with location information stored in <listPlace>.

```
<standOff>
<!-- ========= events ========= -->
  <listEvent>
    <event  xml:id="DB-3F"  where="#WWIIEast"
      when="1942-01">
      <head>Luftwaffe shoots down Ilyushin Il-4</head>
      <head  type="short">DB-3F</head>
      <ab  type="altitude">
        <measure  unit="m"  quantity="7000"/>
      </ab>
      <ab  type="survivor">
        <persName  ref="#p-IC">Иван Чиссов</persName>
      </ab>
    </event>
    <event  xml:id="SCP"  where="#l-GSN"
      when="1943-01-03">
      <head>Snap, Crackle, Pop is shot down by German flak</head>
```

```
    <head type="short">B-17</head>
    <ab type="altitude">
      <measure unit="m" quantity="6700"/>
    </ab>
    <ab type="survivor">
      <persName ref="#p-AM">Alan Magee</persName>
    </ab>
  </event>
  <event xml:id="DS664" where="#l-Obe"
    when="1944-03-24">
    <head>Avro Lancaster shot down by Junkers Ju 88</head>
    <head type="short">Lancaster</head>
    <ab type="altitude">
      <measure unit="m" quantity="5490"/>
    </ab>
    <ab type="survivor">
      <persName ref="#p-NA">Nicholas Alkemade</persName>
    </ab>
  </event>
  <event xml:id="LANSA508" where="#l-PI"
    when="1971-12-04">
    <head>LANSA 508 struck by lightning</head>
    <head type="short">Amazon rainforest</head>
    <ab type="altitude">
      <measure unit="m" quantity="3000"/>
    </ab>
    <ab type="survivor">
      <persName ref="#p-JK">Juliane Diller (née Koepcke)</persName>
    </ab>
  </event>
  <event xml:id="jat367" where="#l-Seb"
    when="1972-01-26T16:01:29Z">
    <head>Bombing of Yugoslav Airlines Flight 367</head>
    <head type="short">JAT 367</head>
    <ab type="altitude">
      <measure unit="m" quantity="10050"/>
    </ab>
    <ab type="survivor">
      <persName ref="#p-VV">Весна Вуловић</persName>
    </ab>
    <ptr type="furtherInfo"
      target="https://en.wikipedia.org/wiki/JAT_Flight_367"/>
  </event>
  <event xml:id="BG" where="#l-KorZ"
    when="1996">
    <head>SAS parachute failure</head>
    <head type="short">Bear Grylls</head>
    <ab type="altitude">
      <measure unit="m" quantity="4877"/>
    </ab>
    <ab type="survivor">
      <persName ref="#p-BG"/>
    </ab>
  </event>
  <event xml:id="CM" where="#l-Joh"
    when="2004-08-22">
    <head>Johannesburg parachute failure</head>
```

```
      <head type="short">Christine McKenzie</head>
      <ab type="altitude">
        <measure unit="m" quantity="3353"/>
      </ab>
      <ab type="survivor">
        <persName ref="#p-CM">Christine McKenzie</persName>
      </ab>
    </event>
    <event xml:id="JBK" where="#l-Kam"
      when="2009-05">
      <head>Cameraman fails to open chute</head>
      <head type="short">Sport videographer</head>
      <ab type="altitude">
        <measure unit="m" quantity="1829"/>
      </ab>
      <ab type="survivor">
        <persName ref="#p-JB"/>
      </ab>
    </event>
  </listEvent>
<!-- ========= places ========= -->
  <listPlace>
    <place xml:id="l-GSN">
      <placeName xml:lang="fr">Gare de Saint-Nazaire</placeName>
      <location>
        <geo>47.28657 -2.21171</geo>
      </location>
    </place>
    <place xml:id="l-Joh">
      <placeName>Johannesburg</placeName>
      <location>
        <country>South Africa</country>
        <region type="province">Gauteng</region>
      </location>
    </place>
    <place xml:id="l-Kam">
      <placeName xml:lang="ru">Камчатке</placeName>
      <location>
        <country>Russia</country>
        <region>Kamchatka Krai</region>
      </location>
    </place>
    <place xml:id="l-Ken">
      <country>Kenya</country>
    </place>
    <place xml:id="l-Obe" xml:lang="de">
      <placeName>Oberkirchen</placeName>
      <location>
        <country>Deutschland</country>
        <region type="state">Nordrhein-Westfalen</region>
        <region>Arnsberg</region>
        <district>Hochsauerlandkreis</district>
        <settlement>Schmallenberg</settlement>
        <geo cert="low">51.154 8.357</geo>
      </location>
    </place>
    <place xml:id="l-PI" xml:lang="es">
```

```
      <placeName>Puerto Inca</placeName>
      <location>
        <country>Peru</country>
        <region>Huánuco</region>
        <district>Puerto Inca</district>
      </location>
    </place>
    <place xml:id="l-Sch">
      <placeName xml:lang="de">Schmallenberg</placeName>
      <placeName xml:lang="wep">Smalmereg</placeName>
      <location/>
    </place>
    <place xml:id="l-Seb">
      <placeName xml:lang="de">Sebnitz</placeName>
      <location>
        <geo>50.966667 14.283333</geo>
      </location>
    </place>
    <place xml:id="l-WWIIEast">
      <name>Eastern Front</name>
      <ptr type="furtherInfo"
        target="https://en.wikipedia.org/wiki/Eastern_Front_(World_War_II)"/>
    </place>
    <place xml:id="l-Zam">
      <country>Zambia</country>
    </place>
  </listPlace>
<!-- ========= people ========= -->
  <listPerson>
    <person xml:id="p-VV">
      <persName xml:lang="sr-Cyrl">Весна Вуловић</persName>
      <persName xml:lang="sr-Latn">Vesna Vulović</persName>
      <birth when="1950-01-03"/>
      <death when="2016-12-23"/>
      <ptr type="furtherInfo"
        target="https://en.wikipedia.org/wiki/Vesna_Vulovi%C4%87"/>
    </person>
    <person xml:id="p-IC">
      <persName xml:lang="uk-Cyrl">Іван Михайлович Чиссов</persName>
      <persName xml:lang="ru-Cyrl">Иван Михайлович Чисов</persName>
      <persName xml:lang="uk-Latn">Ivan Mikhailovich Chisov</persName>
      <birth when="1916"/>
      <death when="1986"/>
      <ptr type="furtherInfo"
        target="https://en.wikipedia.org/wiki/Ivan_Chisov"/>
    </person>
    <person xml:id="p-AM">
      <persName>Alan Eugene Magee</persName>
      <birth when="1919-01-13"/>
      <death when="2003-12-20"/>
      <ptr type="furtherInfo"
        target="https://en.wikipedia.org/wiki/Alan_Magee"/>
    </person>
    <person xml:id="p-NA">
      <persName>Nicholas Stephen Alkemade</persName>
      <birth when="1922-12-10"/>
      <death when="1987-06-22"/>
```

```
      <ptr type="furtherInfo"
        target="https://en.wikipedia.org/wiki/Nicholas_Alkemade"/>
    </person>
    <person xml:id="p-BG">
      <persName>Edward Michael Grylls</persName>
      <birth when="1974-06-07"/>
      <ptr type="furtherInfo"
        target="https://www.beargrylls.com/pages/about-bear-grylls"/>
    </person>
    <person xml:id="p-JK">
      <persName>Juliane Koepcke</persName>
      <birth when="1954-10-10"/>
      <ptr type="furtherInfo"
        target="https://en.wikipedia.org/wiki/Juliane_Koepcke"/>
    </person>
    <person xml:id="p-JB">
      <persName>James Boole</persName>
      <birth>1978, give or take a year</birth>
      <ptr type="furtherInfo"
        target="https://www.facebook.com/james.boole/?ref=page_internal"/>
    </person>
  </listPerson>
<!-- ========= conjunctions and disjunctions ========= -->
  <altGrp type="location">
    <alt xml:id="l-KorZ"
      target="#l-Ken #l-Zam"/>
  </altGrp>
</standOff>
```

The *where* attribute on <event> can point to multiple canonical place descriptions, which indicates that an event took place in multiple locations. For example, the November 2015 Paris Attacks took place at six different locations in Paris; this might be encoded as an <event> with a *where* attribute with six space-separated values, or using a <listPlace> (nested into the <event> or standalone).

Events may also be modeled as hierarchies of nested <event>s. For example, a two part meeting might be represented as follows:

```
<event type="meeting" when="2007-05-29">
  <desc>All day meeting to resolve content models</desc>
  <event type="preamble" notAfter="13:00:00">
    <desc>first part</desc>
  </event>
  <event type="conclusions"
    notBefore="13:00:00">
    <desc>second part</desc>
  </event>
</event>
```

For <event> as a child element of <person>, note the Section *14.3.2.2. Personal Events*.

### 14.3.6 Objects

`<object>` contains a description of a single identifiable physical object.

`<objectName>` (name of an object) contains a proper noun or noun phrase used to refer to an object.

`<listObject>` (list of objects) contains a list of descriptions, each of which provides information about an identifiable physical object.

An object is any material thing whether real, in existence, fictional, missing, or purported about which more information is known. Where objects have proper names the <objectName> element may be used to encode these. However, many objects are not named but the <object> element may still be used to provide a description of them. The <object> element is a more general descriptive form of the <msDesc> element. The latter should be used for describing manuscripts and similar text-bearing objects but can be viewed as a more specific form of the <object> element.

*Please note:* The <object> element is a recent addition to TEI P5 Guidelines as of version 3.5.0 and as such might be more prone to further revision in the next few releases as its use develops. This may be particularly evident where its contents have been borrowed from <msDesc> and have yet to be generalized from their use in the context of manuscript descriptions.

The <object> element usually appears inside the <listObject> element which is used to group descriptions of identifiable objects. The <listObject> element is a member of `model.listLike` and so may appear inside <standOff>, or anywhere else that <list> is allowed. This enables the flexibility of using <listObject> to contain a set of metadata descriptions stored in the TEI header, or as a list of objects transcribed from a source document. The equivalent list for manuscript descriptions is <listBibl>.

`<objectIdentifier>` (object identifier) groups one or more identifiers or pieces of locating information concerning a single object.

`<msContents>` (manuscript contents) describes the intellectual content of a manuscript, manuscript part, or other object either as a series of paragraphs or as a series of structured manuscript items.

`<physDesc>` (physical description) contains a full physical description of a manuscript, manuscript part, or other object optionally subdivided using more specialized elements from the `model.physDescPart` class.

`<history>` (history) groups elements describing the full history of a manuscript, manuscript part, or other object.

`<additional>` (additional) groups additional information, combining bibliographic information about a manuscript or other object, or surrogate copies of it, with curatorial or administrative information.

Overall, the basic structure of an <object> element is akin to that of <msDesc> in that it is providing a structured description of an object. After a group of identifying information, it has the option of paragraphs or, if the `msdescription` module is loaded, <msContents>, <physDesc>, <history>, and <additional> elements for descriptive metadata about this object. Although these elements originate from manuscript description the <object> element may be used for all forms of object (whether text-bearing or not). Where descendents of <object> still have the hallmarks of their use in manuscript description, the descriptions as relating to manuscripts should be interpreted as applying to all forms of object (text-bearing) or not.

The <objectIdentifier> element is a general-purpose grouping element for location or identification information relating to a single object or resource. It is very similar to an <msIdentifier> element with less contraints on the order of its contents. The <objectIdentifier> may be more or less detailed dependent on the needs of the encoder. In some cases an <object> may be used mostly as a common reference point for multiple <objectName> elements to refer back to. In situations, one might provide more detailed information in the <objectIdentifier> where it is available or desirable. Compare

```
<object xml:id="Excalibur-shortIdentifier">
  <objectIdentifier>
    <objectName>Excalibur</objectName>
  </objectIdentifier>
  <p>Excalibur is the name for the legendary sword of King Arthur.</p>
</object>
```

where only a single <objectName> is provided and below where multiple versions are provided.

```
<object xml:id="Excalibur-longerIdentifier">
  <objectIdentifier>
```

```
    <objectName type="main">Excalibur</objectName>
    <objectName type="alt">Caliburn</objectName>
    <objectName xml:lang="cy">Caledfwlch</objectName>
    <objectName xml:lang="cnx">Calesvol</objectName>
    <objectName xml:lang="br">Kaledvoulc'h</objectName>
    <objectName xml:lang="la">Caliburnus</objectName>
    <country>Wales</country>
  </objectIdentifier>
  <p>Excalibur is the main English name for the legendary
     sword of King Arthur. In Welsh it is called
  <mentioned>Caledfwlch</mentioned>, in Cornish it is called
  <mentioned>Calesvol</mentioned>, in Breton it is called
  <mentioned>Kaledvoulc'h</mentioned>, and in Latin it is
     called <mentioned>Caliburnus</mentioned>. In some versions
     of the legend, Excalibur's blade was engraved with phrases on opposite
     sides: <q>Take me up</q> and <q>Cast me away</q> (or similar).</p>
</object>
```

Moreover, the <objectIdentifier> may include an <address> element to provide the address at which the object currently resides. The use of <location> within this enables the provision of geographical coordinates when describing objects not housed in traditional repositories or institutions. This may also be used to supplement more traditional repository location information if available and, for example, to enable providing outputs such as maps showing the location of encoded objects.

```
<objectIdentifier>
  <objectName xml:lang="en">Mask of Tutankhamun</objectName>
  <idno type="carter">256a</idno>
  <idno type="JournalD'Entrée">60672</idno>
  <idno type="exhibition">220</idno>
  <institution>Museum of Egyptian Antiquities</institution>
  <address>
    <street>15 Meret Basha</street>
    <district>Ismailia</district>
    <settlement>Cairo</settlement>
    <country>Egypt</country>
    <location>
      <geo>30.047778 31.233333</geo>
    </location>
  </address>
</objectIdentifier>
```

The <msContents> element is currently used to provide a description of the intellectual contents of any text on an object and, being optional, is not necessary if there are no intellectual contents to describe. (Such contents, especially in the case of artistic objects, may not always be textual.) The <physDesc> element may be used to give a physical description of the object either in prose or using more structured elements as and where they apply to that kind of object. The <history> element provides the option to describe the history of the object as paragraphs or with more structure using the <origin> element, as many <provenance> stages as are appropriate, and <acquisition> to describe its current ownership. The <additional> element may be used to provide information about surrogates for the object (such as digital facsimiles) as well as administrative and curatorial information. A full description of an object can provide more or less detail at any level to represent the state of knowledge about the object.

```
<listObject>
  <object xml:id="TutankhamunMask">
```

```
   <objectIdentifier>
     <objectName  xml:lang="en">Mask of Tutankhamun</objectName>
     <idno  type="carter">256a</idno>
     <idno  type="JournalD'Entrée">60672</idno>
     <idno  type="exhibition">220</idno>
     <institution>Museum of Egyptian Antiquities</institution>
     <address>
       <street>15 Meret Basha</street>
       <district>Ismailia</district>
       <settlement>Cairo</settlement>
       <country>Egypt</country>
       <location>
         <geo>30.047778 31.233333</geo>
       </location>
     </address>
   </objectIdentifier>
   <msContents>
     <p>The back and shoulders of the mask is inscribed with a protective spell in
Egyptian hieroglyphs formed of ten
           vertical and horizontal lines. This spell first appeared on masks in the
Middle Kingdom at least 500 years
           before Tutankhamun, and comes from chapter 151 of the <title>Book of the
Dead</title>.</p>
   </msContents>
   <physDesc>
     <p>The mask of Tutankhamun is 54cm x 39.3cm x 49cm. It is constructed from two
layers of high-karat gold that
           varies in thickness from 1.5-3mm. It weighs approximately 10.23kg and x-ray
crystallography shows that it is
           composed of two alloys of gold with a lighter 18.4 karat shade being used
for the face and neck while a heavier
           22.5 karat gold was used for the rest of the mask.</p>
     <p>In the mask Tutankhamun wears a nemes headcloth which has the royal insignia
of a cobra (Wadjet) and vulture
           (Nekhbet) on it. These are thought respectively to symbolize Tutankhamun's
rule of both Lower Egypt and Upper
           Egypt. His ears are pierced for earrings. The mask has rich inlays of
coloured glass and gemstones, including
           lapis lazuli surrounding the eye and eyebrows, quartz for the eyes,
obsidian for the pupils. The broad collar is
           made up of carnelian, feldspar, turquoise, amazonite, faience and other
stones.</p>
   </physDesc>
   <history>
     <origin>
       <p>The mask of Tutankhamun was created in <origPlace>Egypt</origPlace> around
<origDate  when="-1323"  type="circa">1323 BC</origDate>. It is a death mask of the
18th-dynasty ancient Egyptian Pharaoh Tutankhamun
             who reigned 1332–1323 BC.</p>
     </origin>
     <provenance>
       <p>The mask of Tutankhamun was found in his burial chamber at Theban Necropolis
in the Valley of the Kings in
             1922. On 28 October 1925 the excavation team led by English
archaeologist Howard Carter opened the heavy
             sarcophagus and three coffins and were the first people in around 3,250
years to see the mask of Tutankhamun.
```

```
              Carter wrote in his diary: <quote> The pins removed, the lid was raised.
The penultimate scene was disclosed —
                  a very neatly wrapped mummy of the young king, with golden mask of
sad but tranquil expression, symbolizing
                  Osiris … the mask bears that god's attributes, but the likeness is
that of Tut.Ankh.Amen — placid and
                  beautiful, with the same features as we find upon his statues and
coffins. The mask has fallen slightly
                  back, thus its gaze is straight up to the heavens. </quote>
       </p>
     </provenance>
     <acquisition> In December 1925, the mask was removed from the tomb, placed in a
crate and transported 635
           kilometres (395 mi) to the Egyptian Museum in Cairo, where it remains on
public display. </acquisition>
   </history>
   <additional>
     <adminInfo>
       <custodialHist>
         <custEvent  when="1944">When it was discovered in 1925, the 2.5kg narrow gold
beard was no longer attached to
                  the mask and was reattached to the chin by use of a wooden dowel in
1944.</custEvent>
         <custEvent  when="2014-08"> In August 2014 when the mask was removed from its
display case for cleaning, the
                  beard fell off again. Those working in the museum unadvisedly used a
quick-drying epoxy to attempt to fix
                  it, but left the beard off-centre. </custEvent>
         <custEvent  when="2015-01">The damage was noticed and repaired in January 2015
by a German-Egyptian team who
                  used beeswax, a material known to be used as adhesives by the ancient
Egyptians.</custEvent>
       </custodialHist>
     </adminInfo>
   </additional>
  </object>
</listObject>
```

If the object is being referenced from elsewhere in the document, this is usually done with an <objectName>. For example here the Alfred-Jewel *xml:id* is referenced from a paragraph elsewhere in the document using the *ref* attribute on the <objectName> element.

```
<!-- Inside <standOff>: --><listObject>
  <object  xml:id="Alfred-Jewel">
    <objectIdentifier>
      <country>United Kingdom</country>
      <region>Oxfordshire</region>
      <settlement>Oxford</settlement>
      <institution>University of Oxford</institution>
      <repository>Ashmolean Museum</repository>
      <collection>English Treasures</collection>
      <idno  type="ashmolean">AN1836p.135.371</idno>
      <idno  type="wikipedia">https://en.wikipedia.org/wiki/Alfred_Jewel</idno>
      <objectName>Alfred Jewel</objectName>
    </objectIdentifier>
```

```
   <physDesc>
     <p>The Alfred Jewel is about 6.4 cm in length and is made of combination of
filigreed <material>gold</material>
           surrounding a polished teardrop shaped piece of transparent
<material>quartz</material>. Underneath the rock
           crystal is a cloisonné enamel image of a man with ecclesiastical symbols.
The sides of the jewel holding the
           crystal in place contain an openwork inscription saying "AELFRED MEC HEHT
GEWYRCAN", meaning 'Alfred ordered me
           made'.</p>
   </physDesc>
   <history>
     <origin>It is generally accepted that the Alfred Jewel dates from the
<origDate>late 9th Century</origDate> and
           was most likely made in <origPlace>England</origPlace>. </origin>
     <provenance  when="1693">The jewel was discovered in 1693 at Petherton Park, North
Petherton in the English county
           of Somerset, on land owned by Sir Thomas Wroth. North Petherton is about 8
miles away from Athelney, where King
           Alfred founded a monastery. </provenance>
     <provenance  when="1698">A description of the Alfred Jewel was first published in
1698, in the Philosophical
           Transactions of the Royal Society.</provenance>
     <acquisition> It was bequeathed to Oxford University by Colonel Nathaniel Palmer
(c. 1661-1718) and today is in
           the Ashmolean Museum in Oxford. </acquisition>
   </history>
  </object>
</listObject>
<!-- Inside <text>: -->
<p> The <objectName  ref="#MinsterLovellJewel">Minster Lovell Jewel</objectName> is
probably the most similar to the
<objectName  ref="#Alfred-Jewel">Alfred Jewel</objectName> and was found in
<placeName  ref="#MinsterLovell">Minster
    Lovell</placeName> in <placeName  ref="#Oxfordshire">Oxfordshire</placeName> and
is kept at the <orgName  ref="#AshmoleanMuseum">Ashmolean Museum</orgName>.
</p>
```

There is no restriction on the form, size, or type of object that may be described by an <object> element, however, some objects may be more adequately described by a <place> element depending on context. Where a description of an object is being provided in terms of identification, physical characteristics, or history, then an <object> element may be preferred. Where metadata is being recorded about the geo-political location, population, or similar traits, then the <place> element may be better suited. A corresponding relation between an object description and place may be recorded through the use of the *corresp* attribute. An example of a large object that might be described with the <object> element could be a building such as the Central Library of the National Autonomous University of Mexico.

```
<object  type="building"  xml:lang="en"
  xml:id="UNAM-CL">
  <objectIdentifier>
    <objectName  type="abbr">The Central Library of UNAM</objectName>
    <objectName  type="full">The Central Library of the National Autonomous University
of Mexico</objectName>
    <objectName  xml:lang="es">La Biblioteca Central de la Universidad Nacional
Autónoma de México</objectName>
```

```
   <settlement>Mexico City</settlement>
   <region>Coyoacán</region>
   <country>Mexico</country>
  </objectIdentifier>
  <physDesc>
   <objectDesc>
    <p>The Central Library encompasses an area of <dim  unit="m"  quantity="16000"
       type="area">16 thousand square
           meters</dim> and is built on a three meter platform. The base contains
two <material>basalt</material>
      <objectType>fountains</objectType> and <objectType>decorative
reliefs</objectType> inspired by
          pre-Hispanic art.</p>
    <p>The library has ten windowless floors for book storage, each having enough
space for 120 thousand volumes.
          These storage areas have the necessary lighting, temperature and humidty
conditions for book conservation.
          In the reading room, flanked by a garden on each side, the diffuse and
matte light is filtered through
    <material>thin tecali stone slabs</material>. The semi-basement of the building
contains the service
          and administrative offices of the library.</p>
    <p>The building facades are covered with one of the largest murals in the world
and is made from naturally
          colored stone tiles. It is entitled <title>Historical Representation of
Culture</title> and is by
    <persName  role="artist">Juan O'Gorman</persName>.</p>
   </objectDesc>
   <decoDesc>
    <decoNote>
     <label>Base</label> At the base of the building there are two basalt fountains
and decorative reliefs
          around the outside that are inspired by pre-Hispanic art. The color of the
stone in these elements is left
          exposed to take advantage of the stone's texture as an aesthetic and
expressive element, and to give a
          sense of continuity to the external pavement.</decoNote>
    <decoNote>
     <label>Murals</label>
     <p>The outside windowless portion of the building contains one of the largest
murals in the world. This is
             called <title>Historical Representation of the Culture</title> and is a
<material>stone polychromatic
                mosaic</material> based on the combination of 12 basic colors. The
mural is created in an
             impressionist style where the coloured tiles when seen from a distance
form specific figures. The 12
             colors where chosen from 150 samples of original stones with the
criteria including the stone's
             resistance to degredation by weather. According to the artist,
<persName  role="artist">Juan
                O'Gorman</persName>, in the mosaic he represented three fundamental
historical facets of the Mexican
             culture: the pre-Hispanic era, the Spanish colonial era, and the modern
age. For example with the North
             Wall, this represents the pre-Hispanic era and is dominated by mythical
elements relating to the
```

```
              life-death duality. The left side of the main axis there are
              deities and scenes pertaining to the creation of life. The right hand
side of the mural contains figures
              relating to death. For a more detailed description see
<ref  target="https://en.wikipedia.org/wiki/Central_Library_(UNAM)#Murals">https://en.wikipedia.o
     </decoNote>
   </decoDesc>
  </physDesc>
  <history>
    <origin> In <origDate  when="1948"  type="conceptual">1948</origDate> the architect
and artist Juan O'Gorman, in
        collaboration with architects Gustavo Saavedra and Juan Martinez de Velasco
designed the building with a
        functionalist approach, as part of the greater project of the construction of
the University City on the
        grounds of the Pedregal de San Angel in Mexico City. Originally the building
was planned to host the National
        Library and National Newspaper Library of Mexico. The library finally opened
its doors for the first time on
    <origDate  when="1956-04-05"
       type="opening">5 April 1956</origDate>. In July 2007 it was declared a UNESCO
        world heritage site. </origin>
  </history>
  <additional>
    <adminInfo>
      <custodialHist>
        <custEvent  from="1981"  to="1983"> The library was significantly remodelled
from 1981 - 1983 with the
              purpose of changing from closed shelving to open stacks, providing users
more direct access to the
              collections. </custEvent>
      </custodialHist>
    </adminInfo>
  </additional>
</object>
```

### 14.3.7 Names and Nyms

So far we have discussed ways in which a name or referring string encountered in running text may be resolved by considering the object that the name refers to: in the case of a personal name, the name refers to a person; in the case of a place name, to a place, for example. The resolution of this reference is effected by means of the *key* or *ref* attributes available to all elements which are members of the `att.naming` class, such as <persName> or <placeName> and their more specialized variants such as <forename> or <country>. However, *names* can also be regarded as objects in their own right, irrespective of the objects to which they are attached, notably in onomastic studies. From this point of view, the names *John* in English, *Jean* in French, and *Ivan* in Russian might all be regarded as existing independently of any person to which they are attached, and also independently of any variant forms that might be attested in different sources (such as Jon or Johnny in English, or Jehan or Jojo in French). We use the term *nym* to refer to the canonical or normalized form of a name regarded in such a way, and provide the following elements to encode it:

`<listNym>` (list of canonical names) contains a list of nyms, that is, standardized names for any thing.

`<nym>` (canonical name) contains the definition for a canonical name or name component of any kind.

Any element which is a member of the `att.naming` class may use the attribute *nymRef* to indicate the nym with which it corresponds. Thus, given the following <nym> for the name *Antony*:

```
<listNym>
  <nym  xml:id="N123">
    <form>Antony</form>
  </nym>
<!-- other nym definitions here -->
</listNym>
```

an occurrence of this name in running text might be encoded as follows:

```
<forename  nymRef="#N123">Tony</forename> Blair
```

Note that this association (between ”Tony” and ”Antony”) has nothing to do with any individual who might use the name.

The person identified by this particular Tony may however be indicated independently using the *ref* attribute, either on the forename or on the whole name component:

```
<forename  nymRef="#N123"  ref="#BLT">Tony</forename>
<!-- ... -->
<person  xml:id="BLT">
  <persName>Tony Blair</persName>
  <occupation>politician</occupation>
</person>
```

The <nym> element may be thought of as providing a specialized kind of dictionary entry. Like a dictionary entry, it may contain any element from the `model.entryPart` class, such as <form>, <etym>, etc. For example, we may show that the canonical form for a given nym has two orthographic variants in this way:

```
<nym  xml:id="J451">
  <form>
    <orth  xml:lang="en-US">Ian</orth>
    <orth  xml:lang="en-x-Scots">Iain</orth>
  </form>
</nym>
```

Because a schema intending to make use of the <nym> or <listNym> element must include the `dictionaries` module as well as the `namesdates` module, many other elements are available in addition to <form>. For example, to provide a more complex etymological decomposition of a name, we might use the existing <etym> element, as follows:

```
<nym  xml:id="XYZ">
  <form>Bogomil</form>
  <etym>Means <gloss>favoured by God</gloss> from the <lang>Slavic</lang> elements
<mentioned  xml:lang="ru">bog</mentioned>
    <gloss>God</gloss> and <mentioned  xml:lang="ru">mil</mentioned>
    <gloss>favour</gloss>
  </etym>
</nym>
```

Where it is necessary to mark the substructure of nyms, this may be done by <seg> elements within the <form>:

```
<nym xml:id="ABC">
  <form>
    <choice>
      <seg type="morph">
        <seg>Bog</seg>
        <seg>o</seg>
        <seg>mil</seg>
      </seg>
      <seg type="morph">
        <seg>Bogo</seg>
        <seg>mil</seg>
      </seg>
    </choice>
  </form>
</nym>
```

The <seg> element used here is provided by the TEI `linking` module, which would therefore also need to be included in a schema built to validate such markup. Other possibilities for more detailed linguistic analysis are provided by elements included in that and the `analysis` (see *18. Simple Analytic Mechanisms*) or `iso-fs` modules (see *19. Feature Structures*).

Alternatively, each of the constituents of *Bogomil* might be regarded as a nym in its own right:

```
<nym xml:id="B1" type="part">
  <form>bog</form>
</nym>
<nym xml:id="M1" type="part">
  <form>mil</form>
</nym>
```

Within running text, a name can specify all the nyms associated with it:

```
 ...<name nymRef="#B1 #M1">Bogomil</name>...
```

Similarly, within a nym, the attribute *parts* is used to indicate its constituent parts, where these have been identified as distinct nyms:

```
<nym xml:id="BM1" parts="#B1 #M1">
  <form>Bogomil</form>
</nym>
```

The <nym> element may also combine a number of other <nym> elements together, where it is intended to show that they are all regarded as variations on the same root. Thus the different forms of the name John, all being derived from the same root, may be represented as a hierarchic structure like this:

```
<nym xml:id="J45">
  <form xml:lang="la">Iohannes</form>
  <nym xml:id="J450">
    <form xml:lang="en">John</form>
    <nym xml:id="J4501">
      <form>Johnny</form>
```

```
    </nym>
    <nym xml:id="J4502">
      <form>Jon</form>
    </nym>
  </nym>
  <nym xml:id="J455">
    <form xml:lang="ru">Ivan</form>
  </nym>
  <nym xml:id="J453">
    <form xml:lang="fr">Jean</form>
  </nym>
</nym>
```

The <nym> element may be used for components of geographical or organizational names as well. For example:

```
<geogName ref="tag:projectname.org,2012:LAEI1"
  type="hill">
  <geogFeat xml:lang="gd" nymRef="#LAIRG">Lairig</geogFeat>
  <name>Eilde</name>
</geogName> ... <nym xml:id="LAIRG">
  <form xml:lang="gd">lairig</form>
  <def>sloping hill face</def>
</nym> ...
```

As noted above, use of these elements implies that both the `dictionaries` and the `namesdates` modules are included in a schema.

## 14.4 Dates

The following elements for the encoding of dates and times were introduced in section *3.6.4. Dates and Times*:

`<date>` (date) contains a date in any format.

`<time>` (time) contains a phrase defining a time of day in any format.

The current module `namesdates` provides a mechanism for more detailed encoding of relative dates and times. A *relative temporal expression* describes a date or time with reference to some other (absolute) temporal expression, and thus may contain an <offset> element in addition to one or more <date> or <time> elements:

`<offset>` (offset) marks that part of a relative temporal or spatial expression which indicates the direction of the offset between the two place names, dates, or times involved in the expression.

As members of the `att.datable` and `att.duration` classes, which in turn are members of `att.datable.w3c` and `att.duration.w3c` respectively, the <date> and <time> elements share the following attributes:

`att.datable.w3c` provides attributes for normalization of elements that contain datable events conforming to the W3C XML Schema Part 2: Datatypes Second Edition.

`@when` supplies the value of the date or time in a standard form, e.g. yyyy-mm-dd.

`att.duration.w3c` provides attributes for recording normalized temporal durations.

`@dur` (duration) indicates the length of this element in time.

### 14.4.1 Relative Dates and Times

As noted above, relative dates and times such as ‘in the Two Hundredth and First Year of the Republic’, ‘twenty minutes before noon’, and, more ambiguously, ‘after the lamented death of the Doctor’ or ‘an hour after the game’ have two distinct components. As well as the absolute temporal expression or event to which reference is made (e.g. ‘noon’, ‘the game’, ‘the death of the Doctor’, ‘[the foundation of] the Republic’), they also contain a description of the ‘distance’ between the time or date which is indicated and the referent expression (e.g. ‘the Two Hundredth and First Year’, ‘twenty minutes’, ‘an hour’); and (optionally) an ‘offset’ describing the direction of the distance between the time or date indicated and the referent expression (e.g. ‘of’ implying after, ‘before’, ‘after’).

The 'distance' component of a relative temporal expression may be encoded as a temporal element in its own right using either <date> or <time>, or with the more generic <measure> element. A special element, <offset>, is provided by this module for encoding the 'offset' component of a relative temporal expression. The absolute temporal expression contained within the relative expression may be encoded with a <date> or <time> element; in turn, those elements may of course be relative, and thus contain <date> or <time> elements within themselves. This allows for deeply nested structures such as 'the third Sunday after the first Monday before Lammastide in the fifth year of the King's second marriage ...' but so does natural language.

In the following examples, the *when* and *dur* attributes have been used to simplify processing of variant forms of expression:

```
<date when="1786-12-11">
  <date dur="P14D">A fortnight</date>
  <offset>before</offset>
  <date when="1786-12-25" type="holiday">Christmas 1786</date>
</date>
```

```
I reached the station <time when="14:15:00">
  <time dur="PT30M0S">precisely half an hour</time>
  <offset>after</offset>
  <time when="13:45:00" type="occasion">the departure of the afternoon train to
Boston</time>
</time>
```

In the following example, a nested <date> element is used to show that 'my birthday' and the cited date are parts of the same temporal expression, and hence to disambiguate the phrase 'A week before my birthday on 9th December':

```
<date when="--12-02">
  <date>A week</date>
  <offset>before</offset>
  <date when="--12-09">
    <date type="occasion">my birthday</date> on <date>9th December</date>
  </date>
</date>
```

The alternative reading of this phrase could be encoded as follows:

```
<date when="--12-09">
  <date>A week</date>
  <offset>before</offset>
  <date type="occasion" when="--12-16">my birthday</date> on <date>9th
December</date>
</date>
```

Where more complex or ambiguous expressions are involved, and where it is desirable to make more explicit the interpretive processes required, the feature structure notation described in chapter *19. Feature Structures* may be used. Consider, for example, the following temporal expression which occurs in the *Scottish Temperance Review* of August 1850, referring to the summer holiday known in Glasgow simply as 'the Fair':

```
Not only is the city, <date ana="#gf50">during the Fair</date>, a horrible
nucleus of immorality and wickedness; it sends our multitudes to pollute and
demoralize the country.
```

For the definition of the *ana* attribute, see chapter *18. Simple Analytic Mechanisms* (in particular *18.2. Global Attributes for Simple Analyses*). It is used here to link the temporal phrase with an interpretation of it. Like most traditional fairs and market days, the Glasgow Fair was established by local custom and could vary from year to year. Consequently, in order to provide such an interpretation, it is necessary to draw upon additional information which may or may not be located in the particular text in question. In this case, it is necessary at least to know the spatial and temporal context (year and place) of the fair referred to. These and other features required for the analysis of this particular temporal expression may be combined together as one feature structure of type date-analysis:

```
<fs xml:id="gf50" type="date-analysis">
  <f name="event">
    <string>the Fair</string>
  </f>
  <f name="place">
    <string>Glasgow</string>
  </f>
  <f name="year">
    <numeric value="1850"/>
  </f>
  <f name="from-value">
    <string>1850-08-08</string>
  </f>
  <f name="to-value">
    <string>1850-09-19</string>
  </f>
</fs>
```

For further discussion of feature structure representation see chapter *19. Feature Structures.*

### 14.4.2 Absolute Dates and Times

The following are examples of absolute temporal expressions.

```
The university's view of American affairs
produced a stinging attack by Edmund Burke in the Commons debate of
<date when="1775-10-26">26 October
1775</date>
```

```
<date when="1993-05-14">Friday, 14 May 1993</date>
```

It may be useful to categorize a temporal expression which is given in terms of a named event, such as a public holiday, or a named time such as ‘tea time’ or ‘matins’:

```
In
New York, <date type="occasion" when="--01-01">New Year's Day</date> is the quietest
of holidays, <date when="--07-04" type="occasion">Independence Day</date> the most
turbulent.
```

Absolute temporal expressions denoting times which are given in terms of seconds, minutes, hours, or of well-defined events (e.g. ‘noon’, ‘sunset’) may similarly be represented using the <time> element.

```
The train leaves for Boston at <time  type="twentyfourHour"  when="13:45:00">13:45</time>
```

```
At <time  type="occasion">sunset</time> we walked to the beach.
```

```
The train leaves for Boston at <time  xml:lang="en-US"  type="descriptive"
  when="13:45:00-05:00"> a quarter of two </time>
```

### 14.4.3 More Expressive Normalizations

The attributes for normalization of dates and times so far described use a standard format defined by XML Schema Part 2: Datatypes Second Edition. This format is widely accepted and has significant software support. It is essentially a profile of ISO 8601 *Data elements and interchange formats — Information interchange — Representation of dates and times*. The full ISO standard provides formats not available in the W3C recommendation, for example, the capability to refer to a date by its ordinal date or week date, or to refer to a century. It also provides ways of indicating duration and range.

When this module is included in a schema, the following additional attributes are provided:

`att.datable.iso` provides attributes for normalization of elements that contain datable events using the ISO 8601:2004 standard.

- `@when-iso` supplies the value of a date or time in a standard form.
- `@notBefore-iso` specifies the earliest possible date for the event in standard form, e.g. yyyy-mm-dd.
- `@notAfter-iso` specifies the latest possible date for the event in standard form, e.g. yyyy-mm-dd.
- `@from-iso` indicates the starting point of the period in standard form.
- `@to-iso` indicates the ending point of the period in standard form.

`att.duration.iso` provides attributes for recording normalized temporal durations.

- `@dur-iso` (duration) indicates the length of this element in time.

These attributes may be used in preference to their W3C equivalent when it is necessary to provide a normalized value in some form not supported by the W3C attributes. For example, a century date in the W3C format must be expressed as a range, using the *from* attribute together with either the *to* attribute or the *dur* (duration) attribute:

```
<date  from="1301"  to="1400">fourteenth century</date>
<date  from="1301"  dur="P100Y">fourteenth century</date>
```

With the attribute *when-iso*, however, it is possible to express the same normalized value in any of the following additional ways:

```
<date  when-iso="13">fourteenth
century</date>
<date  when-iso="1301/1400">fourteenth century</date>
<date  when-iso="1301/P100Y">fourteenth century</date>
```

### 14.4.4 Using Non-Gregorian Calendars

All date-related encoding described above makes use of the Gregorian calendar, on which both the ISO and W3C datetime formats are based. However, historical texts often pre-date the invention of the Gregorian calendar in the

16th century, or its adoption in Europe over the following centuries, and many other calendars are used in texts from other cultures and contexts. Non-Gregorian dates can be encoded using methods described below.

First, a Calendar Description element needs to be supplied in the <teiHeader> as described in *2.4.5. Calendar Description*:

```
<calendarDesc>
  <calendar  xml:id="Julian_England">
    <p>The Julian calendar, as used in England until 1752.</p>
  </calendar>
</calendarDesc>
```

The following attributes can now be used to encode dates using this calendar:

`att.calendarSystem` provides attributes for indicating calendar systems to which a date belongs.

- `@calendar` indicates one or more systems or calendars to which the date represented by the content of this element belongs.

`att.datable.custom` provides attributes for normalization of elements that contain datable events to a custom dating system (i.e. other than the Gregorian used by W3 and ISO).

- `@when-custom` supplies the value of a date or time in some custom standard form.
- `@notBefore-custom` specifies the earliest possible date for the event in some custom standard form.
- `@notAfter-custom` specifies the latest possible date for the event in some custom standard form.
- `@from-custom` indicates the starting point of the period in some custom standard form.
- `@to-custom` indicates the ending point of the period in some custom standard form.
- `@datingMethod` supplies a pointer to a <calendar> element or other means of interpreting the values of the custom dating attributes.

```
<p>The Poole by S. <hi>Giles</hi> Churchyarde was a large water in the yeare
<date  calendar="#Julian_England">1244</date>.</p>
```

Here, the *calendar* attribute points to the <calendar> element in the header which defines and describes the calendar used.

The *calendar* attribute is used to specify the calendar used in the *text content* of the dating element which bears it. For reasons of consistency and computability, it is often useful to complement the content of an element with attributes such as *when*, *notBefore*, *notAfter*, etc. For dates expressed in a non-Gregorian calendar, such normalizations may also be expressed in terms of the non-Gregorian calendar specified by the *datingMethod* attribute. The attributes *when-custom*, *notBefore-custom*, *from-custom*, etc. are provided for this purpose; the *datingMethod* attribute is used to identify the calendar used in the value of these attributes:

```
<head> The Tryumphs of Peace.<lb/> That Celebrated the Solemnity of the right
Honorable Sr Francis Iones Knight, at
his Inauguration into the Maioraltie of London, on Monday being the
<date  when-custom="1620-10-30"
   datingMethod="#Julian_England"  calendar="#Julian_England"> 30. of October, 1620.
</date>
</head>
```

Here, the *calendar* attribute specifies the calendar used in the text content of the <date> element, as before, whereas the *datingMethod* attribute signifies that the calendar used in the *when-custom* attribute is also Julian. The schema could be customized in order to constrain the content of custom attributes in a manner similar to the constraints provided on regular Gregorian dating attributes such as *when*, to enforce consistency in the use of non-Gregorian dates.

Custom dating attributes can be combined with any of the standard dating attributes in order to provide a standardized Gregorian version of a non-Gregorian date. We might enhance the preceding example with the addition of *when*, providing the Gregorian calendar equivalent of the Julian date:

```
<date  when-custom="1620-10-30"
  when="1620-11-09"  datingMethod="#Julian_England"
  calendar="#Julian_England"> 30. of
October, 1620. </date>
```

## 14.5 Module for Names and Dates

The module described in this chapter makes available the following components:

**Module namesdates: Names and dates**

- *Elements defined*: addName affiliation age birth bloc climate country death district education event event-Name faith floruit forename genName gender geo geogFeat geogName langKnowledge langKnown listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink nationality nym object objec-tIdentifier objectName occupation offset org orgName persName persPronouns person personGrp persona place placeName population region relation residence roleName settlement sex socecStatus state surname terrain trait
- *Classes defined*: att.datable.custom att.datable.iso model.persNamePart

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 15 Tables, Formulæ, Graphics, and Notated Music

Many documents, both historical and contemporary, include not only text, but also graphics, artwork, and other images. Although some types of images can be represented directly with markup, it is more common practice to include such information by using a reference to an external entity (typically a URL) encoded in a suitable graphical notation.

In addition to graphic images, documents often contain material presented in graphical or tabular format. In such materials, details of layout and presentation may also be of comparatively greater significance or complexity than they are for running text. Indeed, it may often be difficult to make a clear distinction between details relating purely to the rendition of information and those relating to the information itself.

Documents may also contain mathematical formulæ or expressions in other formulaic notations, for which no notation is defined in these Guidelines.

Finally, documents may contain musical notation, embedded in a manner similar to tables, graphs, and formulæ.

These areas (graphics, tabular material, and mathematical or other formulæ, and music) have in common that they have received considerable attention from many other standards bodies or similar professional groups. In part because of this, they may frequently be most conveniently encoded and processed using some notation not defined by these Guidelines. For these reasons, and others, we consider tables, formulæ, graphics, and notated music together in this chapter.

As with text markup in general, many incompatible formats have been proposed for the representation of graphics, formulæ, and tables in electronic form. Unfortunately, no single format as effective as XML in the domain of text has yet emerged for their interchange, to some extent because of the difficulty of representing the information these data formats convey independently of the way it is rendered.

The module defined by this chapter defines special purpose 'container' elements that can be used to encapsulate occurrences of such data within a TEI-conformant document in a portable way. Specific recommendations for the encoding of tables are provided in section *15.1. Tables*, recommendations for mathematical or other formulæ in section *15.2. Formulæ and Mathematical Expressions*, and for the encoding of musical notation in section *15.3. Notated Music in Written Text*. Specific recommendations for the encoding of graphic figures may be found in section *15.4. Specific Elements for Graphic Images*. The rest of the chapter is devoted to general problems of encoding graphic information.

There is at the time of writing no consensus on formats for graphical images, and such formats vary in many ways. We therefore provide (in section *15.5. Overview of Basic Graphics Concepts*) a brief discussion of the ways in which images may be represented, and (in section *15.6. Graphic Image Formats*) a list of formal names for those representations most popular at this time. Each one includes a very brief description. These Guidelines recommend a few particular representations as being the most widely supported and understood.

## 15.1 Tables

A table is the least 'graphic' of the elements discussed in this chapter. Almost any text structure can be presented as a series of rows and columns: one might, for example, choose to show a glossary or other form of list in tabular form, without necessarily regarding it as a table. In such cases, the global *rend* attribute is an appropriate way of indicating that some element is being presented in tabular format; similarly, the global *style* attribute coud be used to provide an appropriate display property in CSS. When tabular presentation is regarded as of less intrinsic importance, it is correspondingly simpler to encode descriptive or functional information about the contents of the table, for example to identify one cell as containing a name and another as containing a date, though the two methods may be combined.

When, however, particular elements are required to encode the tabular arrangement itself, then one or other of the various 'table schemas' now available may be preferable. The schemas in common use generally view a table as a special text element, made up of row elements, themselves composed of cells. Table cells generally appear in row-major order, with the first row from left to right, then the second row, and so on. Details of appearance such as column widths, border lines, and alignment are generally encoded by numerous attributes. Beyond this, however, such schemas differ greatly. This section begins by describing a table schema of this kind; a brief summary of some other widely available table schemas is also provided in section *15.1.2. Other Table Schemas*.

### 15.1.1 TEI Tables

For encoding tables of low to moderate complexity, these Guidelines provide the following special purpose elements:

`<table>` (table) contains text displayed in tabular form, in rows and columns.

`@rows` (rows) indicates the number of rows in the table.

`@cols` (columns) indicates the number of columns in each row of the table.

`<row>` (row) contains one row of a table.

`<cell>` (cell) contains one cell of a table.

The <table> element is defined as a member of the class `model.inter`; it may therefore appear both within other components (such as paragraphs), or between them, provided that the module defined in this chapter has been enabled, as described at the beginning of this chapter.

It is to a large extent arbitrary whether a table should be regarded as a series of rows or as a series of columns. For compatibility with currently available systems, however, these Guidelines require a row-by-row description of a table. It is also possible to describe a table simply as a series of cells; this may be useful for tabular material which is not presented as a simple matrix.

The attributes *rows* and *cols* may be used to indicate the size of a table, or to indicate that a particular cell or row of a table spans more than one row or column. For both tables and cells, rows and columns are always given in top-to-bottom, left-to-right order, although formatting properties such as those provided by CSS may be used to specify that they should be displayed differently. These Guidelines do not require that the size of a table be specified; for most formatting and many other applications, it will be necessary to process the whole table in two passes in any case.

Where cells span more than one column or row, the encoder must determine whether this is a purely presentational effect (in which case the *rend* attribute may be more appropriate), whether the part of the table affected would be better treated as a nested table, or whether to use the spanning attributes listed above.

The *role* attribute may be used to categorize a single cell, or set a default for all the cells in a given row. The present Guidelines distinguish the roles of *label* and *data* only, but the encoder may define other roles, such as 'derived', 'numeric', etc., as appropriate.

These three attributes are provided by the attribute class `att.tableDecoration` of which both <cell> and <row> are members; see further *1.3.1. Attribute Classes*.

The following simple example demonstrates how the data presented as a labelled list in section *3.8. Lists* might be represented by an encoder wishing to preserve its original appearance as a table:

```
<table rend="boxed" rows="2" cols="2">
  <head rend="it">Report of the conduct and progress of Ernest
     Pontifex. Upper Vth form — half term ending Midsummer 1851</head>
  <row>
    <cell role="label">Classics</cell>
    <cell>Idle listless and unimproving</cell>
  </row>
  <row>
    <cell role="label">Mathematics</cell>
    <cell>ditto</cell>
  </row>
  <row>
    <cell role="label">Divinity</cell>
    <cell>ditto</cell>
  </row>
  <row>
    <cell role="label">Conduct in house</cell>
    <cell>Orderly</cell>
  </row>
  <row>
    <cell role="label">General conduct</cell>
    <cell>Not satisfactory, on account of his great unpunctuality and
        inattention to duties</cell>
  </row>
</table>
```

Note that this encoding makes no attempt to represent the full significance of the 'ditto' cells above; these might be regarded as simple links between the cells containing them and that to which they refer, or as virtual copies of it. For ways of representing either interpretation, see chapter *17. Linking, Segmentation, and Alignment*.

The following example demonstrates how a simple statistical table may be represented using this scheme:

```
<table rows="4" cols="4">
  <head>Poor Man's Lodgings in Norfolk (Mayhew, 1843)</head>
  <row role="label">
    <cell/>
    <cell>Dossing Cribs or Lodging Houses</cell>
    <cell>Beds</cell>
    <cell>Needys or Nightly Lodgers</cell>
  </row>
  <row>
    <cell role="label">Bury St Edmund's</cell>
    <cell>5</cell>
    <cell>8</cell>
    <cell>128</cell>
  </row>
  <row>
    <cell role="label">Thetford</cell>
    <cell>3</cell>
    <cell>6</cell>
    <cell>36</cell>
  </row>
  <row>
    <cell role="label">Attleboro'</cell>
    <cell>3</cell>
    <cell>5</cell>
    <cell>20</cell>
  </row>
  <row>
    <cell role="label">Wymondham</cell>
    <cell>1</cell>
    <cell>11</cell>
    <cell>22</cell>
  </row>
</table>
```

Note the use of a blank cell in the first row to ensure that the column labels are correctly aligned with the data. Again, this encoding does not explicitly represent the alignment between column and row labels and the data to which they apply. Where the primary emphasis of an encoding is on the semantic content of a table, a more explicit mechanism for the representation of structured information such as that provided by the feature structure mechanism described in chapter *19. Feature Structures* may be preferred. Alternatively, the general purpose linkage and alignment mechanisms described in chapter *17. Linking, Segmentation, and Alignment* may also be applied to individual cells of a table.

The content of a table cell need not be simply character data. It may also contain any sequence of the phrase-level elements described in chapter *3. Elements Available in All TEI Documents*, thus allowing for the encoding of potentially more useful semantic information, as in the following example, where the fact that one cell contains a number and the other contains a place name has been explicitly recorded:

```
<table>
  <head>US State populations, 1990</head>
  <row>
    <cell>
```

```
      <name>Wyoming</name>
    </cell>
    <cell>
      <num>453,588</num>
    </cell>
  </row>
  <row>
    <cell>
      <name>Alaska</name>
    </cell>
    <cell>
      <num>550,043</num>
    </cell>
  </row>
  <row>
    <cell>
      <name>Montana</name>
    </cell>
    <cell>
      <num>799,065</num>
    </cell>
  </row>
  <row>
    <cell>
      <name>Rhode Island</name>
    </cell>
    <cell>
      <num>1,003,464</num>
    </cell>
  </row>
</table>
```

The use of semantically marked elements within a <cell> enables the encoder to convey something about the nature and significance of the information, rather than merely suggesting how to display it in rows and columns.

Alternatively, the *role* attribute might be used to convey such information:

```
<table>
  <head>US State populations, 1990</head>
  <row>
    <cell  role="statename">Wyoming </cell>
    <cell  role="pop">453,588 </cell>
  </row>
  <row>
    <cell  role="statename">Alaska </cell>
    <cell  role="pop">550,043 </cell>
  </row>
  <row>
    <cell  role="statename">Montana </cell>
    <cell  role="pop">799,065 </cell>
  </row>
  <row>
    <cell  role="statename">Rhode Island</cell>
    <cell  role="pop">1,003,464</cell>
  </row>
</table>
```

The content of table elements is not limited to <head> and <row>. Milestone elements such as <cb> and <lb> allow new columns or lines to be signalled inside tables; <figure> provides an option for including data which is not amenable to normal row and cell analysis; and other elements such as <epigraph> and <trailer> provide options for including text which is clearly part of the table, but outside the actual tabular layout. This example shows the use of <trailer>:

```
<table>
  <head>The Table of Battallions, reduced out of the grand square of
     men.</head>
  <row>
    <cell>1</cell>
    <cell>2</cell>
    <cell>3</cell>
    <cell>4</cell>
  </row>
<!-- ... -->
  <row>
    <cell>841</cell>
    <cell>3</cell>
    <cell>289 289 256</cell>
    <cell>7</cell>
  </row>
  <trailer>The end of the Table of Battallions reduced out of the
     battels of g. and squares of men: vpon the right side of euery
     leafe.</trailer>
</table>
```

### 15.1.2 Other Table Schemas

Many authoring systems include built-in support for their own or for public table schemas. These provide an enhanced user interface and good formatting capabilities, but are often product-specific, despite their use of a XML markup language.

The DTD developed by the Association of American Publishers (AAP) and standardized in ANSI Z39.59 provided a very simple encoding for correspondingly simple tables. This has been further developed, together with the table DTD documented in ISO Technical Report 9537, and now forms part of ISO 12083. The TEI table model described above has functionality very similar to that defined by ISO 12083.

For more complex tables, the most effective publicly-available DTD is probably that developed by the US Department of Defense CALS project. This supports vertical and horizontal spanning and various kinds of text rotation and justification within cells and is also directly supported by a number of existing XML software systems.

The CALS table model is much too complex to describe fully here; for historical background see http://archive.is/gGzsZ; for more recent simplifications of it and current implementations see https://www.oasis-open.org/specs/tablemodels.php. As with any other XML vocabulary, the XML version of the CALS model may readily be included in a TEI schema, using the techniques described in *24.3. Customization*.

The XHTML table model (?) is based on the HTML table model (?). Both models support arrangement of arbitrary data into rows and columns of cells. Table rows and columns may be grouped to convey additional structural information and may be rendered by user agents in ways that emphasize this structure. Support for incremental rendering of tables and for rendering on 'non-visual' user agents is also available. Special elements and attributes are provided to associate metadata with tables. They indicate the table's purpose, or are for the benefit of people using speech or Braille-based user agents. Tables are not recommended for use purely as a means to lay out document content, as this leads to many accessibility problems (see further https://www.w3.org/TR/WCAG10-HTML-TECHS/#tables). Stylesheets provide a far more effective means of controlling layout and other visual characteristics in both HTML and XML documents.

## 15.2 Formulæ and Mathematical Expressions

Mathematical and chemical formulæ pose problems similar to those posed by tables in that rendition may be of great significance and hard to disentangle from content. They also require access to a wide range of special characters,

for most of which standard entity names already exist in the documented ISO entity sets (see further chapters *vi Languages and Character Sets* and *5. Characters, Glyphs, and Writing Modes*).

Formulæ and tables are also similar in that well-researched and detailed DTD fragments have already been developed for them independently of the TEI. They differ in that (for mathematics at least) there also exists a richly detailed text-based but non-XML notation which is very widely used: this is the TeX system, and the sets of descriptive macros developed for it such as LaTeX, AMS-TeX, and AMS-LaTeX.

The AAP and ISO standards mentioned in section *15.1. Tables* above both provide DTDs for equations as well as for tables, which now form part of ISO 12083. The European Mathematical Trust, an organization set up specifically to enhance research support for European mathematicians, has also defined a general purpose mathematical DTD known as EuroMath (`http://xml.coverpages.org/emt-ukc-index.html`), for which it provides both software and services.

Most if not all of the functionality provided by these DTDs can now be found in the OpenMath and MathML XML-based systems briefly described below.

As with tables, in all the XML solutions a tension exists between the need to encode the way a formula is written (its appearance) and the need to represent its semantics. If the object of the encoding is purely to act as an interchange format among different formatting programs, then there is no need to represent the mathematical meaning of an expression. If however the object is to use the encoding as input to an algebraic manipulation system (such as Mathematica or Maple) or a database system, clearly simply representing superscripts and subscripts will be inadequate.

The <formula> element provided by these Guidelines makes no attempt to represent the internal structure of formulæ.

`<formula>` (formula) contains a mathematical or other formula.

By default, a <formula> is assumed to contain character data which is not validated in any way. The notation used may however be named, using the *notation* attribute provided by the `att.notated` class.

```
<formula notation="TeX">$e=mc^2$</formula>
```

The character data must still be well-formed, of course, which means that < and *&* must be escaped with entity references or numeric character references, e.g.

```
<formula notation="TeX">$\matrix{0 &amp;
1\cr&lt;0&amp;>1}$</formula>
```

Alternatively, if more detailed markup is desired, the content of the <formula> element may be redefined to include elements defined by some other XML vocabulary, such as that of ISO 12083, or to use elements from the OpenMath or MathML schemas.

When the content of a <formula> element is not expressed in XML the notation used should always be specified using the *notation* attribute as above, and in the following longer example:

```
<p>Achilles runs ten times faster than the tortoise and gives the
animal a headstart of ten meters. Achilles runs those ten meters, the
tortoise one; Achilles runs that meter, the tortoise runs a decimeter;
Achilles runs that decimeter, the tortoise runs a centimeter; Achilles
runs that centimeter, the tortoise, a millimeter; Fleet-footed
Achilles, the millimeter, the tortoise, a tenth of a millimeter, and
so on to infinity, without the tortoise ever being overtaken. . . Such
is the customary version.
<!-- ... --> The problem does not change, as
you can see; but I would like to know the name of the poet who
provided it with a hero and a tortoise. To those magical competitors
and to the series <formula notation="TeX">$$ {1 \over 10} + {1 \over
    100} + {1 \over 1000} + {1 \over 10,\!000} + \dots $$</formula> the
argument owes its fame.</p>
```

The *notation* attribute supplies the name of a notation ('TeX'), which is expected to be identified somewhere in document metadata.

Mathematical Markup Language (MathML) (?) is a vocabulary for describing mathematical notation, capturing both its structure and content. It provides two types of markup: Presentation Markup, which captures the *notational* structure of an expression and could be seen as the 'TeX for the Web' and Content Markup, which captures the *mathematical* structure of an expression. Most of its content elements correspond with the range of operators, relations, and named functions typically found at the high-school level of mathematics. The tortoise example given above in TeX can be re-expressed in MathML as

```
<m:math>
  <m:mfrac>
    <m:mrow>
      <m:mn>1</m:mn>
    </m:mrow>
    <m:mrow>
      <m:mn>10</m:mn>
    </m:mrow>
  </m:mfrac>
  <m:mo>+</m:mo>
  <m:mfrac>
    <m:mrow>
      <m:mn>1</m:mn>
    </m:mrow>
    <m:mrow>
      <m:mn>100</m:mn>
    </m:mrow>
  </m:mfrac>
  <m:mo>+</m:mo>
  <m:mfrac>
    <m:mrow>
      <m:mn>1</m:mn>
    </m:mrow>
    <m:mrow>
      <m:mn>1000</m:mn>
    </m:mrow>
  </m:mfrac>
  <m:mo>+</m:mo>
  <m:mfrac>
    <m:mrow>
      <m:mn>1</m:mn>
    </m:mrow>
    <m:mrow>
      <m:mn>10000</m:mn>
    </m:mrow>
  </m:mfrac>
  <m:mo>+</m:mo>
  <m:mo>…</m:mo>
</m:math>
```

MathML 2.0 provides support for a 'Semantic Math-Web', XML namespaces, and other current XML standards, such as XML DOM, OMG IDL, ECMAScript, and Java. It also provides a modularized version of the MathML DTD so that MathML fragments 'embedded' in XHTML 1.1 documents can be correctly validated.

The OpenMath (`https://openmath.org/standard/`) project is coordinated by the OpenMath Society (`https://openmath.org/`) and funded by the European Commission under the Esprit Multimedia Standards Initiative that commenced in September 1997. It is likely to become a key standard for communicating semantically rich representations of mathematical objects both on and off the Web in a platform-independent manner.

The OpenMath Standard (https://openmath.org/standard/om20-2004-06-30/) consists of specifications for

1. OpenMath objects, representing the structure of formulæ (https://openmath.org/standard/om20-2004-06-30/omstd20html-2.xml#cha_obj);
2. Content Dictionaries, providing semantic context (https://openmath.org/standard/om20-2004-06-30/omstd20html-4.xml#cha_cd);
3. Encodings, both binary (https://openmath.org/standard/om20-2004-06-30/omstd20html-3.xml#sec_binary) and XML (https://openmath.org/standard/om20-2004-06-30/omstd20html-3.xml#sec_xml).

OpenMath and MathML have certain common aspects. They both use prefix operators, both are XML-based and they both construct their objects by applying certain rules recursively. Such similarities facilitate mapping between the two standards. There are also some key differences between MathML and OpenMath. OpenMath does not provide support for presentation of mathematical objects and its scope of semantically-oriented elements is much broader that of MathML, with the expressive power to cover virtually all areas of computational mathematics. In fact, a particular set of Content Dictionaries, the 'MathML CD Group', covers the same areas of mathematics as the Content Markup elements of MathML 2.0.

Finally, OMDoc (http://omdoc.org/) is an extension of the OpenMath standard that supplies markup for structures such as axioms, theorems, proofs, definitions, texts (mixing formal content with mathematical text).

In-line versus block placement for an equation can be distinguished if desired, via the global *rend* attribute. The global *n* and *xml:id* attributes may also be used to label or identify the formula, as in the following example:

```
<p>The volume of a sphere
is given by the formula: <formula  xml:id="f12"  n="12"  rend="inline">
    <m:math>
      <m:mi>V</m:mi>
      <m:mo>=</m:mo>
      <m:mfrac>
        <m:mrow>
          <m:mn>4</m:mn>
        </m:mrow>
        <m:mrow>
          <m:mn>3</m:mn>
        </m:mrow>
      </m:mfrac>
      <m:mi>π</m:mi>
      <m:msup>
        <m:mrow>
          <m:mi>r</m:mi>
        </m:mrow>
        <m:mrow>
          <m:mn>3</m:mn>
        </m:mrow>
      </m:msup>
    </m:math>
  </formula> which is readily calculated.</p>
<p>As we have seen in equation <ptr  target="#f12"/>, ... </p>
```

## 15.3 Notated Music in Written Text

Music, like many other art forms, is often mentioned, discussed and described in writings of various kinds. This applies to both historical and contemporary documents, even though methods of notating music have changed considerably in western history. In most cases, music notation enters the text flow in a way similar to figures, images, or graphs. On other occasions, elements of music notation are treated as inline characters in running text.

<notatedMusic> provides a way to signal the presence of music notation in text, but defer to other representations, which are not covered by the TEI guidelines, to describe the music notation itself. In fact several commercial, academic, and standard bodies have developed digital representations of music notation. Given the topic's complexity, these representations often focus on different aspects and adopt different methodologies. Therefore, <notatedMusic> only defines a container element to encode the occurrence of music notation and allows linking to the data format preferred by the encoder. (Note: <notatedMusic> is not the same as <musicNotation>, a metadata element, which is used to describe musical notation that appears in a manuscript. See *11. Manuscript Description*.)

The following elements can be used for encoding music notation in text:

`<notatedMusic>` encodes the presence of music notation in a text.

`<ptr>` (pointer) defines a pointer to another location.

`<desc>` (description) contains a short description of the purpose, function, or use of its parent element, or when the parent is a documentation element, describes or defines the object being documented.

`<graphic>` (graphic) indicates the location of a graphic or illustration, either forming part of a text, or providing an image of it.

`<binaryObject>` provides encoded binary data representing an inline graphic, audio, video or other object.

groups elements representing or containing music notation.

can be used to indicate the location of a representation of the music notation.

- *mimeType* supplies the MIME type of the data format, when available.

can be used to give a prose description of the notated music.

can be used to indicate the location of a graphical representation of the music notation.

provides encoded binary data which constitutes another representation of the music notation (e.g. audio).

The <notatedMusic> element may contain a textual description and pointers to various representations of the music notation in different media. An external representation of the notated music is specified using the <ptr> element, whose *target* attribute provides its electronically-accessible location. The attribute *mimeType* supplies the MIME type of the data format when available. For example:

```
<notatedMusic>
  <ptr  target="bar1.xml"/>
</notatedMusic>
```

A textual description of the notation can be provided within the <desc> element; alternatively, a <label> may be supplied. For example:

```
<notatedMusic>
  <ptr  target="bar1.xml"/>
  <desc>First bar of Chopin's Scherzo No.3 Op.39</desc>
</notatedMusic>
```

It is possible to link to any kind of music notation data format. However, when a MIME type is not available, it is recommended that the format be specified in the description. See the following examples.

MIME type available:

```
<notatedMusic>
  <ptr  target="bar1.xml"
    mimeType="application/vnd.recordare.musicxml"/>
  <desc>First bar of Chopin's Scherzo No.3 Op.39. Encoded in
     MusicXML.</desc>
</notatedMusic>
```

MIME type not available:

```
<notatedMusic>
  <ptr  target="bar1.ly"/>
  <desc>First bar of Chopin's Scherzo No.3 Op.39. Encoded in
     Lilypond.</desc>
</notatedMusic>
```

Application format:

```
<notatedMusic>
  <ptr  target="bar1.mscz"
    mimeType="application/x-musescore"/>
  <desc>First bar of Chopin's Scherzo No.3 Op.39. MuseScore Notation
     Software format.</desc>
</notatedMusic>
```

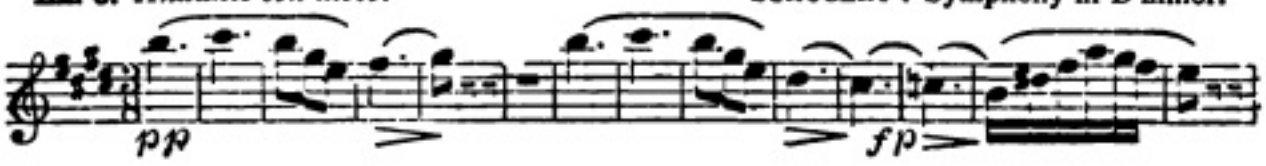
It is possible to specify the location of digital objects representing the notated music in other media such as images or audio-visual files. The interpretation of the correspondence between the notated music and these digital objects is not encoded explicitly. We recommend the use of <graphic> and <binaryObject> mainly as a fallback mechanism when the notated music format is not displayable by the application using the encoding. The alignment of encoded notated music, images carrying the notation, and audio files is a complex matter for which we refer the reader to other formats and specifications such as MPEG-SMR.

```
<notatedMusic>
  <ptr  target="bar1.xml"/>
  <graphic  url="bar1.jpg"/>
  <desc>First bar of Chopin's Scherzo No.3 Op.39</desc>
</notatedMusic>
```

It is also recommended, when useful, to embed XML-based music notation formats, such as the Music Encoding Initiative format as content of <notatedMusic>. This must be done by means of customization.

In modern printing, music notation positioned between blocks of text for illustrative purposes is usually referred to as a 'figure' or 'example'. In this cases, we recommend the inclusion of <notatedMusic> in <figure> in order to encode possible captions and headers. For example:

67. We now give some examples, from the works of the great masters, of some of the most frequently used bowings.

Ex. 3. *Andante con moto.* SCHUBERT: Symphony in B minor.

```
<div  n="67">
  <p> We now give some examples, from the works of the great masters, of
     some of the most frequently used bowings. </p>
  <figure  n="Ex. 3">
```

```
    <head>SCHUBERT: Symphony in B minor.</head>
    <notatedMusic>
      <ptr  target="example_schubert.xml"/>
    </notatedMusic>
  </figure>
</div>
```

## 15.4 Specific Elements for Graphic Images

The following special purpose elements are used to indicate the presence of graphic images within a document:

`<figure>` (figure) groups elements representing or containing graphic information such as an illustration, formula, or figure.

`<graphic>` (graphic) indicates the location of a graphic or illustration, either forming part of a text, or providing an image of it.

`<binaryObject>` provides encoded binary data representing an inline graphic, audio, video or other object.

`<figDesc>` (description of figure) contains a brief prose description of the appearance or content of a graphic figure, for use when documenting an image without displaying it.

The <graphic> and <binaryObject> elements form part of the common core module, and are discussed in section *3.10. Graphics and Other Non-textual Components*.

The <figure> element is used to contain images, captions, and textual descriptions of the pictures. The images themselves are specified using the <graphic> element, whose *url* attribute provides the location of an image. For example:

```
<figure>
  <graphic  url="Fig1.pdf"/>
</figure>
```

Three kinds of content may be supplied inside a <figure> element: the element <head> may be used to transcribe (or supply) a descriptive heading or title for the graphic itself as in this example:

```
<figure>
  <graphic  url="Fig1.pdf"/>
  <head>The View from the Bridge</head>
</figure>
```

Figures are often accompanied not only by a title or heading (a caption), but by a paragraph or so of commentary (a legend) following the caption. One or more <p> or <ab> elements may be used to transcribe any commentary on the figure in the source:

```
<figure>
  <graphic  url="pullman.png"/>
  <head>Above:</head>
  <p>The drawing room of the Pullman house, the white and gold saloon
     where the magnate delighted in giving receptions for several hundred
     people.</p>
  <figDesc>The figure shows an elaborately decorated room, at least
     twenty-five feet side to side and fifty feet long, with ornate
     mouldings and Corinthian columns on the walls, overstuffed armchairs
     and loveseats arranged in several conversational groupings, and two
     large chandeliers.</figDesc>
</figure>
```

Here, the figure contains a heading 'Above' which is complemented by a paragraph of description 'The drawing room … several hundred people'. Both of these are transcribed from the source, while the description is provided by the encoder, for use by applications which cannot display the graphic directly. In documents created in electronic form with the needs of print-handicapped readers in mind, the <figDesc> element may be provided by the author rather than a subsequent encoder.

```
<figure>
  <graphic url="Fig1.jpg"/>
  <head>Figure One: The View from the Bridge</head>
  <figDesc>A Whistleresque view showing four or five sailing boats in
     the foreground, and a series of buoys strung out between
     them.</figDesc>
</figure>
```

Where the graphic itself contains large amounts of text, perhaps with a complex structure, and perhaps difficult to distinguish from the graphic, the encoder should choose whether to regard the graphic as containing the text (in which case, a nested <floatingText> element may be included within the <figure> element) or to regard the enclosed text as being a separate division of the <text> element in which the graphic appears. In this latter case, an appropriate  or <div1> (etc.) element may be used for the text represented within the graphic, and the <figure> element embedded within it. The choice will depend to a large degree on the encoder's understanding of the relationship between the graphic and the surrounding text.

A figure which is internally divided, or contains sub-figures, may be encoded with nested <figure> elements, as in the following example.

```
<figure n="6.45">
  <figure n="a">
    <graphic url="./figs/6.45a.png"/>
    <ab type="caption">Parallel</ab>
  </figure>
  <figure n="b">
    <graphic url="./figs/6.45b.png"/>
    <ab type="caption">Perspective</ab>
  </figure>
  <ab type="caption">The two canonical view volumes, for the (a)
     parallel and (b) perspective projections. Note that -z is to the
     right.</ab>
</figure>
```

Like any other element in the TEI scheme, figures may be given identifiers so that they can be aligned with other elements, and linked to or from them, as described in chapter *17. Linking, Segmentation, and Alignment*. Some common examples are discussed briefly here; full information is provided in that chapter.

It is often desirable to maintain two versions of an image in an electronic file: one a low resolution or 'thumbnail' version which, when selected by the user, causes the other, high resolution, version to be accessed. In TEI terms, the thumbnail image acts as a *reference* to the other. Supposing that a thumbnail version of the figure discussed above is available as `fig1th.png`, we might embed a reference to the image using the simple <ref> element discussed in section *3.7. Simple Links and Cross-References*:

```
<ref target="#IM1">Click
here <graphic url="fig1th.png"/> for enlightenment </ref>
<figure xml:id="IM1">
  <graphic url="fig1.jpg"/>
</figure>
```

Another common requirement is to associate part or the whole of an image with a textual element not necessarily contiguous to it in the text; this is sometimes known as a *callout*. When the module for transcription is included in a schema, specific attributes for parts of a text and parts (or all) of a digital image are available; these are discussed in *12.1. Digital Facsimiles*. In addition, chapter *17. Linking, Segmentation, and Alignment* may be consulted for other mechanisms available for this purpose.

The following example assumes that we wish to associate one portion of the image held as 'fig1' with chapter two of some text, and another portion of it with chapter three. The application may be thought of as a hypertext browser in which the user selects from a graphic image which part of a text to read next, but the mechanism is independent of this particular application.

The first requirement is some way of identifying and hence pointing to sub-parts of a graphic image. This may be done by pointing into an XML graphic representation, for example an SVG file. Thus

```
<ptr xml:id="PD1" target="Fig1.svg#object1"/>
<ptr xml:id="PD2" target="Fig1.svg#object2"/>
```

These <ptr> elements identify two areas within the image 'Fig1' by pointing at elements inside the XML file `Fig1.svg`, which contains the following.

```
<svg xmlns="http://www.w3.org/2000/svg" width="8cm" height="3cm"
 viewBox="2 1 8 3">
 <g id="object1">
  <ellipse style="fill: #ffffff" cx="3.875"
   cy="3.025" rx="1.175" ry="1.175"/>
 </g>
 <g id="object2">
  <rect style="fill: #a81616" x="7.8"
   y="1.9" width="2.17581" height="2.24833"/>
 </g>
</svg>
```

The next requirement is some way of identifying the parts of the document to which a link is to be made. The most obvious way of doing this is to use the global *xml:id* attribute:

```
<div1 type="chapter" xml:id="CHAP1">
<!-- ... -->
</div1>
<div1 type="chapter" xml:id="CHAP2">
<!-- ... -->
</div1>
```

Now, all that is needed to linking these areas to the relevant chapters is a <linkGrp> element, as described in section *17.1. Links*:

```
<linkGrp type="callout">
 <link target="#CHAP1 #PD1"/>
 <link target="#CHAP2 #PD2"/>
</linkGrp>
```

In this example, the SVG representation of the graphic is stored externally to the TEI document and linked by means of a pointer. It is also possible to embed the SVG representation directly within the TEI by extending the content model of the <figure> element to permit an element `<svg>` from the SVG namespace. Like other customizations of the TEI scheme, this is carried out using the techniques documented in section *1.2. Defining a TEI Schema*; further examples are provided in chapter *17. Linking, Segmentation, and Alignment*.

## 15.5 Overview of Basic Graphics Concepts

The first major distinction in graphic representation is that between raster graphics and vector graphics. A *raster image* is a list of points, or dots. Scanners, fax machines, and other simple devices easily produce digital raster images, and such images are therefore quite common. A *vector image*, in contrast, is a list of geometrical objects, such as lines, circles, arcs, or even cubes. These are much more difficult to produce, and so are mainly encountered as the output of sophisticated systems such as architectural and engineering CAD programs.

Raster images are difficult to modify because by definition they only encode single points: a line, for example, cannot grow or shrink as such, since it is not identified as such. Only its component parts are identified, and only they can be manipulated. Therefore the resolution or dot-size of a raster image is important, which is not the case with vector images. It is also far more difficult to convert raster images to vector images than to perform the opposite conversion. Raster images generally require more storage space than vector images, and a wide variety of methods exists for compressing them; the variation in these methods leads to corresponding variations in representations for storage and transmission of raster images.

Motion video usually consists of a long series of raster images. Data compression is even more effective on video than on single raster images (mainly owing to redundancy which arises from the usual similarity of adjacent frames). Notations for representing full-motion video are hotly debated at this time, and any user of these Guidelines would do well to obtain up-to-date expert advice before undertaking a project using them.

The compression methods used with any of these image types may be 'lossy' or 'lossless'. Methods for *lossy compression* save space by discarding a small portion of the image's detail, such as fine distinctions of shading. When decompressed, therefore, such an image will be only a close approximation of the original. In contrast, *lossless compression* guarantees that the exact uncompressed image will be reproducible from the compressed form: only truly redundant information is removed. In general, therefore, lossless compression does not save quite so much space as lossy compression, though it does guarantee fidelity to the original uncompressed image.

Raster images may be characterized by their *resolution*, which is the number of dots per inch used to represent the image. Doubling the resolution will give a more precise image, but also quadruple the storage requirement (before compression), and affect processing time for any operations to be performed, such as displaying an image for a reader. Motion video also has resolution in time: the number of frames to be shown per second. Encoders should consider carefully what resolution(s) and frame rate(s) to use for particular applications; these Guidelines express no recommendation in this matter, save the universal ones of consistency and documentation.

Within any image, it is typical to refer to locations via Cartesian coordinate axes: values for x, y, and sometimes z and/or time. However, graphic notations vary in whether coordinates count from left-to-right and top-to-bottom, or another way. They also vary in whether coordinates are considered real (inches, millimeters, and so on), or virtual (dots). These Guidelines do not recommend any of these methods over another, but all decisions made should be applied consistently, and documented in the <encodingDesc> section of the TEI header.[91]

Methods of aligning images and text are discussed in *12.1. Digital Facsimiles*.

The chromatic values of an image may be rendered in many different ways. In monochrome images every displayed point is either black or white. In *grayscale* images, each point is rendered in some shade of gray, the number of shades varying from system to system. In true polychrome images, points are rendered in different hues, again with varying limitations affecting the number of distinct shades and the means by which they are displayed.

## 15.6 Graphic Image Formats

As noted above, there exists a wide variety of different graphics formats, and the following list is in no way exhaustive. Moreover, inclusion of any format in this list should not be taken as indicating endorsement by the TEI of this format or any products associated with it. Some of the formats listed here are proprietary to a greater or lesser extent and cannot therefore be regarded as standards in any meaningful sense. They are however widely used by many different vendors.

The following formats are widely used at the present time, and likely to remain supported by more than one vendor's software:

- BMP: Microsoft bitmap format
- CGM: Computer Graphics Metafile

[91] Since no special purpose element is provided for this purpose by the current version of these Guidelines, such information should be provided as one or more distinct paragraphs at the end of the <encodingDesc> element described in section *2.3. The Encoding Description*.

- GIF: Graphics Interchange Format
- JPEG: Joint Photographic Expert Group
- PBM: Portable Bit Map
- PCX: IBM PC raster format
- PICT: Macintosh drawing format
- PNG: Portable Network Graphics format
- Photo-CD: Kodak Photo Compact Disk format
- QuickTime: Apple real-time image system
- SMIL: Synchronized Multimedia Integration Language format
- SVG: Scalable Vector Graphics format
- TIFF: Tagged Image File Format

Brief descriptions of all the above are given below. Where possible, current addresses or other contact information are shown for the originator of each format. Many formal standards, especially those promulgated by ISO and many related national organizations (ANSI, DIN, BSI, and many more), are available from those national organizations. Addresses may be found in any standard organizational directory for the country in question.

### 15.6.1 Vector Graphic Formats

**CGM: Computer Graphics Metafile** This vector graphics format is specified by an ISO standard, ISO 8632:1987, amended in 1990. It defines binary, character, and plain-text encodings; the non-binary forms are safer for blind interchange, especially over networks. Documentation on CGM is available from ISO and from its member national bodies such as AFNOR, ANSI, BSI, DIN, JIS, etc.

**SVG: Scalable Vector Graphics format** SVG is a language for describing two-dimensional vector and mixed vector or raster graphics in XML. It is defined by the Scalable Vector Graphics (SVG) 1.0 Specification, W3C Recommendation, 04 September 2001, and is available at `https://www.w3.org/TR/2001/REC-SVG-20010904/`.

**PICT: Macintosh drawing format** This format is universally supported on Macintosh™ systems, and readable by a limited range of software for other systems. Documentation is available from Apple Computer Company, Cupertino, California USA.

### 15.6.2 Raster Graphic Formats

**PNG: Portable Network Graphics format** PNG is a non-proprietary raster format currently widely available. It provides an extensible file format for the lossless, portable, well-compressed storage of raster images. Indexed-color, grayscale, and truecolor images are supported, plus an optional alpha channel. Sample depths range from 1 to 16 bits. It is defined by IETF RFC 2083, March 1997.

**TIFF: Tagged Image File Format** Currently the most widely supported raster image format, especially for black and white images, TIFF is also one of the few formats commonly supported on more than one operating system. The drawback to TIFF is that it actually is a wrapper for several formats, and some TIFF-supporting software does not support all variants. TIFF files may use LZW, CCITT Group 4, or PackBits compression methods, or may use no compression at all. Also, TIFF files may be monochrome, grayscale, or polychromatic. All such options should be specified in prose at the end of the <encodingDesc> section of the TEI header for any document including TIFF images. TIFF is owned by Aldus Corporation. Documentation on TIFF is available from them at Craigcook Castle, Craigcook Road, Edinburgh EH4 3UH, Scotland, or 411 First Avenue South, Seattle, Washington 98104 USA.

**GIF: Graphics Interchange Format** Raster images are widely available in this form, which was created by CompuServe Information Services, but has by now been implemented for many other systems as well. Documentation on GIF is copyright by, and is available from, CompuServe Incorporated, Graphics Technology Department, 5000 Arlington Center Boulevard, Columbus, Ohio 43220 USA.

**PBM: Portable Bit Map** PBM files are easy to process, eschewing all compression in favor of transparency of file format. PBM files can, of course, be compressed by generic file-compression tools for storage and transfer.

Public domain software exists which will convert many other formats to and from PBM. Documentation on PBM is copyright by Jeff Poskanzer, and is available widely on the Internet.

**PCX: IBM PC raster format** This format is used by most IBM PC paint programs, and supports both monochrome and polychromatic images. Documentation is available from ZSoft Corporation, Technical Support Department, ATTN: Technical Reference Manual, 450 Franklin Rd. Suite 100, Marietta, GA 30067 USA.

**BMP: Microsoft bitmap format** This format is the standard raster format for computer using Microsoft Windows™ or Presentation Manager™. Documentation is available from Microsoft Corporation.

### 15.6.3 Photographic and Motion Video Formats

**JPEG: Joint Photographic Experts Group** This standard is sponsored by CCITT and by ISO. It is ISO/IEC Draft International Standard 10918-1, and CCITT T.81. It handles monochrome and polychromatic images with a variety of compression techniques. JPEG per se, like CCITT Group IV, must be encapsulated before transmission; this can be done via TIFF, or via the JPEG File Interchange Format (JFIF), as commonly done for Internet delivery.

**QuickTime: Apple real-time image system** QuickTime is a proprietary method introduced by Apple Computer Company to synchronize the display of various data. The data can include frames of video, sound, lighting control mechanisms, and other things. Viewers for QuickTime productions are available for Apple and other computers. Further information is available from Apple Computer Incorporated, 10201 North de Anza Boulevard MS 23AQ, Cupertino, California 95014 USA.

**Photo-CD: Kodak Photo Compact Disk format** This format was introduced by Kodak for rasterizing photographs and storing them on CD-ROMs (about one hundred 35mm file images fit on one disk), for display on televisions or CD-I systems. Information on Photo-CD is available from Kodak Limited, Research and Development, Headstone Drive, Harrow, Middlesex HA1 4TY, UK.

**SMIL: Synchronized Multimedia Integration Language format** SMIL is a W3C Recommendation which supports the integration of independent multimedia objects into a synchronized multimedia presentation. It provides multimedia authors with easily-defined basic timing relationships, fine-tuned synchronization, spatial layout, direct inclusion of non-text and non-image media objects, hyperlink support for time-based media, and adaptiveness to varying user and system characteristics. SMIL 1.0 (`https://www.w3.org/TR/REC-smil/`) became a W3C Recommendation on June 15, 1998, and was further developed in SMIL 2.0. SMIL 2.0 adds native support for transitions, animation, event-based interaction, extended layout facilities, and more sophisticated timing and synchronization primitives to the SMIL 1.0 language. It also allows reuse of SMIL syntax and semantics in other XML-based languages, in particular those who need to represent timing and synchronization. For example, SMIL 2.0 components are used for integrating timing into XHTML Document Types and into SVG. SMIL 2.0 also provides recommendations for Document Types based on SMIL 2.0 Modules (`https://www.w3.org/TR/2005/REC-SMIL2-20050107/smil-modules.html`). One such Document Type is the SMIL 2.0 Language Profile (`https://www.w3.org/TR/2005/REC-SMIL2-20050107/smil20-profile.html`). It contains support for all of the major SMIL 2.0 features including animation, content control, layout, linking, media object, meta-information, structure, timing, and transition effects and is designed for Web clients that support direct playback from SMIL 2.0 markup. SMIL 2.0 (`https://www.w3.org/TR/2001/REC-smil20-20010807/`) became a W3C Recommendation on August 7, 2001, becoming the first vocabulary to provide XML Schema support and to have reached such status.

As noted above, the reader will encounter many, many other graphics formats.

## 15.7 Module for Tables, Formulæ, Notated Music, and Graphics

The module described in this chapter provides the following features:

**Module figures: Tables, formulæ, notated music, and figures**

- *Elements defined*: cell figDesc figure formula notatedMusic row table
- *Classes defined*: att.tableDecoration

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema.*

# 16 Language Corpora

The term *language corpus* is used to mean a number of rather different things. It may refer simply to any collection of linguistic data (for example, written, spoken, signed, or multimodal), although many practitioners prefer to reserve it for collections which have been organized or collected with a particular end in view, generally to characterize a particular state or variety of one or more languages. Because opinions as to the best method of achieving this goal differ, various subcategories of corpora have also been identified. For our purposes however, the distinguishing characteristic of a corpus is that its components have been selected or structured according to some conscious set of design criteria.

These design criteria may be very simple and undemanding, or very sophisticated. A corpus may be intended to represent (in the statistical sense) a particular linguistic variety or sublanguage, or it may be intended to represent all aspects of some assumed 'core' language. A corpus may be made up of whole texts or of fragments or text samples. It may be a 'closed' corpus, or an 'open' or 'monitor' corpus, the composition of which may change over time. However, since an open corpus is of necessity finite at any particular point in time, the only likely effect of its expansibility from the encoding point of view may be some increased difficulty in maintaining consistent encoding practices (see further section *16.5. Recommendations for the Encoding of Large Corpora*). For simplicity, therefore, our discussion largely concerns ways of encoding closed corpora, regarded as single but composite texts.

Language corpora are regarded by these Guidelines as *composite texts* rather than *unitary texts* (on this distinction, see chapter *4. Default Text Structure*). This is because although each discrete sample of language in a corpus clearly has a claim to be considered as a text in its own right, it is also regarded as a subdivision of some larger object, if only for convenience of analysis. Corpora share a number of characteristics with other types of composite texts, including anthologies and collections. Most notably, different components of composite texts may exhibit different structural properties (for example, some may be composed of verse, and others of prose), thus potentially requiring elements from different TEI modules.

Aside from these high-level structural differences, and possibly differences of scale, the encoding of language corpora and the encoding of individual texts present identical sets of problems. Any of the encoding techniques and elements presented in other chapters of these Guidelines may therefore prove relevant to some aspect of corpus encoding and may be used in corpora. Therefore, we do not repeat here the discussion of such fundamental matters as the representation of multiple character sets (see chapter *vi Languages and Character Sets*); nor do we attempt to summarize the variety of elements provided for encoding basic structural features such as quoted or highlighted phrases, cross-references, lists, notes, editorial changes and reference systems (see chapter *3. Elements Available in All TEI Documents*). In addition to these general purpose elements, these Guidelines offer a range of more specialized sets of tags which may be of use in certain specialized corpora, for example those consisting primarily of verse (chapter *6. Verse*), drama (chapter *7. Performance Texts*), transcriptions of spoken text (chapter *8. Transcriptions of Speech*), etc. Chapter *1. The TEI Infrastructure* should be reviewed for details of how these and other components of these Guidelines should be tailored to create a TEI customization appropriate to a given application. In sum, it should not be assumed that only the matters specifically addressed in this chapter are of importance for corpus creators.

This chapter does however include some other material relevant to corpora and corpus-building, for which no other location appeared suitable. It begins with a review of the distinction between unitary and composite texts, and of the different methods provided by these Guidelines for representing composite texts of different kinds (section *16.1. Varieties of Composite Text*). Section *16.2. Contextual Information* describes a set of additional header elements provided for the documentation of contextual information, of importance largely though not exclusively to language corpora. This is the additional module for language corpora proper. Section *16.3. Associating Contextual Information with a Text* discusses a mechanism by which individual parts of the TEI header may be associated with different parts of a TEI-conformant text. Section *16.4. Linguistic Annotation of Corpora* reviews various methods of providing linguistic annotation in corpora, with some specific examples of relevance to current practice in corpus linguistics. Finally, section *16.5. Recommendations for the Encoding of Large Corpora* provides some general recommendations about the use of these Guidelines in the building of large corpora.

## 16.1 Varieties of Composite Text

Both unitary and composite texts may be encoded using these Guidelines; composite texts, including corpora, will typically make use of the following tags for their top-level organization.

`<teiCorpus>` (TEI corpus) contains the whole of a TEI encoded corpus, comprising a single corpus header and one or more <TEI> elements, each containing a single text header and a text.

`<TEI>` (TEI document) contains a single TEI-conformant document, combining a single TEI header with one or more members of the `model.resource` class. Multiple <TEI> elements may be combined within a <TEI> (or <teiCorpus>) element.

`<teiHeader>` (TEI header) supplies descriptive and declarative metadata associated with a digital resource or set of resources.

`<text>` (text) contains a single text of any kind, whether unitary or composite, for example a poem or drama, a collection of essays, a novel, a dictionary, or a corpus sample.

`<group>` (group) contains the body of a composite text, grouping together a sequence of distinct texts (or groups of such texts) which are regarded as a unit for some purpose, for example the collected works of an author, a sequence of prose essays, etc.

Full descriptions of these may be found in chapter *2. The TEI Header* (for <teiHeader>), and chapter *4. Default Text Structure* (for <teiCorpus>, <TEI>, <text>, and <group>); this section discusses their application to composite texts in particular.

In these Guidelines, the word *text* refers to any stretch of discourse, whether complete or incomplete, unitary or composite, which the encoder chooses (perhaps merely for purposes of analytic convenience) to regard as a unit. The term *composite text* refers to texts within which other texts appear; the following common cases may be distinguished:

- language corpora
- collections or anthologies
- poem cycles and epistolary works (novels or essays written in the form of collections or series of letters)
- otherwise unitary texts, within which one or more subordinate texts are embedded

The elements listed above may be combined to encode each of these varieties of composite text in different ways.

In corpora, the component samples are clearly distinct texts, but the systematic collection, standardized preparation, and common markup of the corpus often make it useful to treat the entire corpus as a unit, too. Some corpora may become so well established as to be regarded as texts in their own right; the Brown and LOB corpora are now close to achieving this status.

The <teiCorpus> element is intended for the encoding of language corpora, though it may also be useful in encoding newspapers, electronic anthologies, and other disparate collections of material. The <TEI> element may be used in the same manner itself; the <teiCorpus> element, however, makes explicit the multiplicity of the collection, whatever it may be. The individual samples in the corpus are encoded as separate <TEI> elements, and the entire corpus is enclosed in a <TEI> or <teiCorpus> element. Each sample has the usual structure for a <TEI> document, comprising a <teiHeader> followed by one or more members of the `model.resource` class. The corpus, too, has a corpus-level <teiHeader> element, in which the corpus as a whole, and encoding practices common to multiple samples may be described. The overall structure of a TEI-conformant corpus is thus:

```
<teiCorpus xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!-- ... -->
  </teiHeader>
  <TEI>
    <teiHeader>
<!-- ... -->
    </teiHeader>
    <text>
<!-- ... -->
    </text>
  </TEI>
  <TEI>
    <teiHeader>
<!-- ... -->
    </teiHeader>
    <text>
```

```
<!-- ... -->
    </text>
  </TEI>
</teiCorpus>
```

Or, alternatively:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!-- ... -->
  </teiHeader>
  <TEI xmlns="http://www.tei-c.org/ns/1.0">
    <teiHeader>
<!-- ... -->
    </teiHeader>
    <text>
<!-- ... -->
    </text>
  </TEI>
  <TEI xmlns="http://www.tei-c.org/ns/1.0">
    <teiHeader>
<!-- ... -->
    </teiHeader>
    <text>
<!-- ... -->
    </text>
  </TEI>
</TEI>
```

Header information which relates to the whole corpus rather than to individual components of it should be factored out and included in the <teiHeader> element prefixed to the whole. This two-level structure allows for contextual information to be specified at the corpus level, at the individual text level, or at both. Discussion of the kinds of information which may thus be specified is provided below, in section *16.2. Contextual Information*, as well as in chapter *2. The TEI Header*. Information of this type should in general be specified only once: a variety of methods are provided for associating it with individual components of a corpus, as further described in section *16.3. Associating Contextual Information with a Text*.

In some cases, the design of a corpus is reflected in its internal structure. For example, a corpus of newspaper extracts might be arranged to combine all stories of one type (reportage, editorial, reviews, etc.) into some higher-level grouping, possibly with sub-groups for date, region, etc. A <teiCorpus> element may occur directly inside a <teiCorpus> specifically to allow direct support for reflecting such internal corpus structure in the markup. For example:

```
<teiCorpus xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!-- metadata for entire corpus -->
  </teiHeader>
  <teiCorpus xmlns="http://www.tei-c.org/ns/1.0">
    <teiHeader>
<!-- metadata for news sub-corpus -->
    </teiHeader>
    <TEI>
      <teiHeader>
<!-- metadata for article #1 -->
```

```
     </teiHeader>
     <text>
<!-- transcription of article #1 -->
     </text>
    </TEI>
    <TEI>
     <teiHeader>
<!-- metadata for article #2 -->
     </teiHeader>
     <text>
<!-- transcription of article #2 -->
     </text>
    </TEI>
    <TEI>
     <teiHeader>
<!-- metadata for article #3 -->
     </teiHeader>
     <text>
<!-- transcription of article #3 -->
     </text>
    </TEI>
    <TEI>
     <teiHeader>
<!-- metadata for article #4 -->
     </teiHeader>
     <text>
<!-- transcription of article #4 -->
     </text>
    </TEI>
  </teiCorpus>
  <teiCorpus xmlns="http://www.tei-c.org/ns/1.0">
    <teiHeader>
<!-- metadata for editorial sub-corpus -->
    </teiHeader>
    <TEI>
     <teiHeader>
<!-- metadata for editorial #1 -->
     </teiHeader>
     <text>
<!-- transcription of editorial #1 -->
     </text>
    </TEI>
    <TEI>
     <teiHeader>
<!-- metadata for editorial #2 -->
     </teiHeader>
     <text>
<!-- transcription of editorial #2 -->
     </text>
    </TEI>
  </teiCorpus>
  <teiCorpus xmlns="http://www.tei-c.org/ns/1.0">
    <teiHeader>
<!-- metadata for reviews -->
    </teiHeader>
    <TEI>
     <teiHeader>
```

```
<!-- metadata for review #1 -->
     </teiHeader>
     <text>
<!-- transcription of review #1 -->
     </text>
   </TEI>
   <TEI>
     <teiHeader>
<!-- metadata for review #2 -->
     </teiHeader>
     <text>
<!-- transcription of review #2 -->
     </text>
   </TEI>
  </teiCorpus>
</teiCorpus>
```

Furthermore, useful groupings of components may easily be expressed using the text classification and identification elements described in section *16.2.1. The Text Description*, and those for associating declarations with corpus components described in section *16.3. Associating Contextual Information with a Text*. These methods also allow several different methods of text grouping to co-exist, each to be used as needed at different times. This helps minimize the danger of cross-classification and misclassification of samples, and helps improve the flexibility with which parts of a corpus may be characterized for different applications.

Anthologies and collections are often treated as texts in their own right, if only for historical reasons. In conventional publishing, at least, anthologies are published as units, with single editorial responsibility and common front and back matter which may need to be included in their electronic encodings. The texts collected in the anthology, of course, may also need to be identifiable as distinct individual objects for study.

Poem cycles, epistolary novels, and epistolary essays differ from anthologies in that they are often written as single works, by single authors, for single occasions; nevertheless, it can be useful to treat their constituent parts as individual texts, as well as the cycle itself. Structurally, therefore, they may be treated in the same way as anthologies: in both cases, the body of the text is composed largely of other texts.

The <group> element is provided to simplify the encoding of collections, anthologies, and cyclic works; the <group> element may also be used to record the potentially complex internal structure of language corpora. (For a full description, see chapter *4. Default Text Structure*.) The choice between using <group> or nested <teiCorpus> elements is up to individual encoders, but in general when it is useful to associate a significant quantity of metadata with such a unit of text it is easier to use <teiCorpus>.

Some composite texts, finally, are neither corpora, nor anthologies, nor cyclic works: they are otherwise unitary texts within which other texts are embedded. In general, they may be treated in the same way as unitary texts, using the normal <TEI> and <body> elements. The embedded text itself may be encoded using the <text> element. For further discussion, see chapter *4. Default Text Structure*.

All composite texts share the characteristic that their different component texts may be of structurally similar or dissimilar types. If all component texts may all be encoded using the same module, then no problem arises. If however they require different modules, then these must be included in the TEI customization. This process is described in more detail in section *1.1. TEI Modules*.

## 16.2 Contextual Information

Contextual information is of particular importance for collections or corpora composed of samples from a variety of different kinds of text. Examples of such contextual information include: the age, sex, and geographical origins of participants in a language interaction, or their socio-economic status; the cost and publication data of a newspaper; the topic, register or factuality of an extract from a textbook. Such information may be of the first importance, whether as an organizing principle in creating a corpus (for example, to ensure that the range of values in such a parameter is evenly represented throughout the corpus, or represented proportionately to the population being sampled), or as a selection criterion in analysing the corpus (for example, to investigate the language usage of some particular vector of social characteristics).

Such contextual information is potentially of equal importance for unitary texts, and these Guidelines accordingly make no particular distinction between the kinds of information which should be gathered for unitary and for composite texts. In either case, the information should be recorded in the appropriate section of a TEI header, as described in chapter *2. The TEI Header*. In the case of language corpora, such information may be gathered together in the overall corpus header, or split across all the component texts of a corpus, in their individual headers, or divided between the two. The association between an individual corpus text and the contextual information applicable to it may be made in a number of ways, as further discussed in section *16.3. Associating Contextual Information with a Text* below.

Chapter *2. The TEI Header*, which should be read in conjunction with the present section, describes in full the range of elements available for the encoding of information relating to the electronic file itself, for example its bibliographic description and those of the source or sources from which it was derived (see section *2.2. The File Description*); information about the encoding practices followed with the corpus, for example its design principles, editorial practices, reference system, etc. (see section *2.3. The Encoding Description*); more detailed descriptive information about the creation and content of the corpus, such as the languages used within it and any descriptive classification system used (see section *2.4. The Profile Description*); and version information documenting any changes made in the electronic text (see section *2.6. The Revision Description*).

In addition to the elements defined by chapter *2. The TEI Header*, several other elements can be used in the TEI header if the additional module defined by this chapter is invoked. These additional tags make it possible to characterize the social or other situation within which a language interaction takes place or is experienced, the physical setting of a language interaction, and the participants in it. Though this information may be relevant to, and provided for, unitary texts as well as for collections or corpora, it is more often recorded for the components of systematically developed corpora than for isolated texts, and thus this module is referred to as being 'for language corpora'.

When the module defined in this chapter is included in a schema, a number of additional elements become available within the <profileDesc> element of the TEI header (discussed in section *2.4. The Profile Description*).

`<textDesc>` (text description) provides a description of a text in terms of its situational parameters.

`<particDesc>` (participation description) describes the identifiable speakers, voices, or other participants in any kind of text or other persons named or otherwise referred to in a text, edition, or metadata.

`<settingDesc>` (setting description) describes the setting or settings within which a language interaction takes place, or other places otherwise referred to in a text, edition, or metadata.

These elements, members of the `model.profileDescPart`, are discussed in the remainder of the chapter.

### 16.2.1 The Text Description

The <textDesc> element provides a full description of the situation within which a text was produced or experienced, and thus characterizes it in a way relatively independent of any *a priori* theory of text-types. It is provided as an alternative or a supplement to the common use of descriptive taxonomies used to categorize texts, which is fully described in section *2.4.3. The Text Classification*, and section *2.3.7. The Classification Declaration*. The description is organized as a set of values and optional prose descriptions for the following eight *situational parameters*, each represented by one of the following eight elements:

`<channel>` (primary channel) describes the medium or channel by which a text is delivered or experienced. For a written text, this might be print, manuscript, email, etc.; for a spoken one, radio, telephone, face-to-face, etc.

`@mode` specifies the mode of this channel with respect to speech and writing.

`<constitution>` (constitution) describes the internal composition of a text or text sample, for example as fragmentary, complete, etc.

`@type` specifies how the text was constituted.

`<derivation>` (derivation) describes the nature and extent of originality of this text.

`@type` categorizes the derivation of the text.

`<domain>` (domain of use) describes the most important social context in which the text was realized or for which it is intended, for example private vs. public, education, religion, etc.

`@type` categorizes the domain of use.

`<factuality>` (factuality) describes the extent to which the text may be regarded as imaginative or non-imaginative, that is, as describing a fictional or a non-fictional world.

`@type` categorizes the factuality of the text.

`<interaction>` (interaction) describes the extent, cardinality and nature of any interaction among those producing and experiencing the text, for example in the form of response or interjection, commentary, etc.

`@type` specifies the degree of interaction between active and passive participants in the text.

`@active` specifies the number of active participants (or *addressors*) producing parts of the text.

`@passive` specifies the number of passive participants (or *addressees*) to whom a text is directed or in whose presence it is created or performed.

`<preparedness>` (preparedness) describes the extent to which a text may be regarded as prepared or spontaneous.

`@type` a keyword characterizing the type of preparedness.

`<purpose>` characterizes a single purpose or communicative function of the text.

`@type` specifies a particular kind of purpose.

`@degree` specifies the extent to which this purpose predominates.

These elements constitute a model class called `model.textDescPart`; new parameters may be defined by defining new elements and adding them to that class, as further described in *24.3. Customization*.

By default, a text description will contain each of the above elements, supplied in the order specified. Except for the <purpose> element, which may be repeated to indicate multiple purposes, no element should appear more than once within a single text description. Each element may be empty, or may contain a brief qualification or more detailed description of the value expressed by its attributes. It should be noted that some texts, in particular literary ones, may resist unambiguous classification in some of these dimensions; in such cases, the situational parameter in question should be given the content 'not applicable' or an equivalent phrase.

Texts may be described along many dimensions, according to many different taxonomies. No generally accepted consensus as to how such taxonomies should be defined has yet emerged, despite the best efforts of many corpus linguists, text linguists, sociolinguists, rhetoricians, and literary theorists over the years. Rather than attempting the task of proposing a single taxonomy of *text-types* (or the equally impossible one of enumerating all those which have been proposed previously), the closed set of *situational parameters* described above can be used in combination to supply useful distinguishing descriptive features of individual texts, without insisting on a system of discrete high-level text-types. Such text-types may however be used in combination with the parameters proposed here, with the advantage that the internal structure of each such text-type can be specified in terms of the parameters proposed. This approach has the following analytical advantages:[92]

- it enables a relatively continuous characterization of texts (in contrast to discrete categories based on type or topic)
- it enables meaningful comparisons across corpora
- it allows analysts to build and compare their own text-types based on the particular parameters of interest to them
- it is equally applicable to spoken, written, or signed texts

Two alternative approaches to the use of these parameters are supported by these Guidelines. One is to use pre-existing taxonomies such as those used in subject classification or other types of text categorization. Such taxonomies may also be appropriate for the description of the topics addressed by particular texts. Elements for this purpose are described in section *2.4.3. The Text Classification*, and elements for defining or declaring such classification schemes in section *2.3.7. The Classification Declaration.* A second approach is to develop an application-specific set of *feature structures* and an associated *feature system declaration,* as described in chapters *19. Feature Structures* and *19.11. Feature System Declaration.*

[92] Schemes similar to that proposed here were developed in the 1960s and 1970s by researchers such as Hymes, Halliday, and Crystal and Davy, but have rarely been implemented; one notable exception being the pioneering work on the Helsinki Diachronic Corpus of English, on which see ?

Where the organizing principles of a corpus or collection so permit, it may be convenient to regard a particular set of values for the situational parameters listed in this section as forming a *text-type* in its own right; this may also be useful where the same set of values applies to several texts within a corpus. In such a case, the set of text-types so defined should be regarded as a *taxonomy*. The mechanisms described in section *2.3.7. The Classification Declaration* may be used to define hierarchic taxonomies of such text-types, provided that the <catDesc> component of the <category> element contains a <textDesc> element rather than a prose description. Particular texts may then be associated with such definitions using the mechanisms described in sections *2.4.3. The Text Classification*.

Using these situational parameters, an informal domestic conversation might be characterized as follows:

```
<textDesc n="Informal domestic conversation">
  <channel mode="s">informal face-to-face conversation</channel>
  <constitution type="single">each text represents a continuously
     recorded interaction among the specified participants
  </constitution>
  <derivation type="original"/>
  <domain type="domestic">plans for coming week, local affairs</domain>
  <factuality type="mixed">mostly factual, some jokes</factuality>
  <interaction type="complete"
    active="plural" passive="many"/>
  <preparedness type="spontaneous"/>
  <purpose type="entertain" degree="high"/>
  <purpose type="inform" degree="medium"/>
</textDesc>
```

The following example demonstrates how the same situational parameters might be used to characterize a novel:

```
<textDesc n="novel">
  <channel mode="w">print; part issues</channel>
  <constitution type="single"/>
  <derivation type="original"/>
  <domain type="art"/>
  <factuality type="fiction"/>
  <interaction type="none"/>
  <preparedness type="prepared"/>
  <purpose type="entertain" degree="high"/>
  <purpose type="inform" degree="medium"/>
</textDesc>
```

### 16.2.2 The Participant Description

The <particDesc> element in the <profileDesc> element provides additional information about the participants in a spoken text or, where this is judged appropriate, the persons named or depicted in a written text. When the detailed elements provided by the `namesdates` module described in *14. Names, Dates, People, and Places* are included in a schema, this element can contain detailed demographic or descriptive information about individual speakers or groups of speakers, such as their names or other personal characteristics. Individually identified persons may also identified by a code which can then be used elsewhere within the encoded text, for example as the value of a *who* attribute.

It should be noted that although the terms *speaker* or *participant* are used throughout this section, it is intended that the same mechanisms may be used to characterize fictional personæ or 'voices' within a written text, except where otherwise stated. For the purposes of analysis of language usage, the information specified here should be equally applicable to written, spoken, or signed texts.

The element <particDesc> contains a description of the participants in an interaction, which may be supplied as straightforward prose, possibly containing a list of names, encoded using the usual <list> and <name> elements, or alternatively using the more specific and detailed <listPerson> element provided by the `namesdates` module described in *14. Names, Dates, People, and Places*.

For example, a participant in a recorded conversation might be described informally as follows:

```
<particDesc xml:id="p2">
  <p>Female informant, well-educated, born in Shropshire UK, 12 Jan
     1950, of unknown occupation. Speaks French fluently.
     Socio-Economic status B2 in the PEP classification scheme.</p>
</particDesc>
```

Alternatively, when the namesdates module is included in a schema, information about the same participant described above might be provided in a more structured way as follows:

```
<person sex="2" age="mid">
  <birth when="1950-01-12">
    <date>12 Jan 1950</date>
    <name type="place">Shropshire, UK</name>
  </birth>
  <langKnowledge tags="en fr">
    <langKnown level="first" tag="en">English</langKnown>
    <langKnown tag="fr">French</langKnown>
  </langKnowledge>
  <residence>Long term resident of Hull</residence>
  <education>University postgraduate</education>
  <occupation>Unknown</occupation>
  <socecStatus scheme="#pep" code="#b2"/>
</person>
```

An identified character in a drama or a novel may also be regarded as a participant in this sense, and encoded using the same techniques:[93]

```
<particDesc>
  <p>The chief speaking characters in this novel are
  <list>
     <item xml:id="EMWOO">
       <name>Emma Woodhouse</name>
     </item>
     <item xml:id="DARCY">
       <name>Mr Darcy</name>
     </item>
<!-- ... -->
    </list>
  </p>
</particDesc>
```

Here, the characters are simply listed without the detailed structure which use of the <listPerson> element permits.

### 16.2.3 The Setting Description

The <settingDesc> element is used to describe the setting or settings in which language interaction takes place. It may contain a prose description, analogous to a stage description at the start of a play, stating in broad terms the locale, or a more detailed description of a series of such settings.

Each distinct setting is described by means of a <setting> element.

`<setting>` describes one particular setting in which a language interaction takes place.

[93] It is particularly useful to define participants in a dramatic text in this way, since it enables the *who* attribute to be used to link <sp> elements to definitions for their speakers; see further section *7.2.2. Speeches and Speakers*.

Individual settings may be associated with particular participants by means of the optional *who* attribute which this element inherits as a member of the `att.ascribed` if, for example, participants are in different places. This attribute identifies one or more individual participants or participant groups, as discussed earlier in section *16.2.2. The Participant Description*. If this attribute is not specified, the setting details provided are assumed to apply to all participants represented in the language interaction. Note however that it is not possible to encode different settings for the same participant: a participant is deemed to be a person within a specific setting.

The <setting> element may contain either a prose description or a selection of elements from the classes `model.nameLike.agent`, `model.dateLike`, or `model.settingPart`. By default, when the module defined by this chapter is included in a schema, these classes thus provide the following elements:

`<name>` (name, proper noun) contains a proper noun or noun phrase.

`<date>` (date) contains a date in any format.

`<time>` (time) contains a phrase defining a time of day in any format.

`<locale>` contains a brief informal description of the kind of place concerned, for example: a room, a restaurant, a park bench, etc.

`<activity>` (activity) contains a brief informal description of what a participant in a language interaction is doing other than speaking, if anything.

Additional more specific naming elements such as <orgName> or <persName> may also be available if the `namesdates` module is also included in the schema.

The following example demonstrates the kind of background information often required to support transcriptions of language interactions, first encoded as a simple prose narrative:

```
<settingDesc>
  <p>The time is early spring, 1989. P1 and P2 are playing on the rug
     of a suburban home in Bedford. P3 is doing the washing up at the
     sink. P4 (a radio announcer) is in a broadcasting studio in
     London.</p>
</settingDesc>
```

The same information might be represented more formally in the following way:

```
<settingDesc>
  <setting  who="#p1 #p2">
    <name  type="city">Bedford</name>
    <name  type="region">UK: South East</name>
    <date>early spring, 1989</date>
    <locale>rug of a suburban home</locale>
    <activity>playing</activity>
  </setting>
  <setting  who="#p3">
    <name  type="city">Bedford</name>
    <name  type="region">UK: South East</name>
    <date>early spring, 1989</date>
    <locale>at the sink</locale>
    <activity>washing-up</activity>
  </setting>
  <setting  who="#p4">
    <name  type="place">London, UK</name>
    <time>unknown</time>
    <locale>broadcasting studio</locale>
    <activity>radio performance</activity>
  </setting>
</settingDesc>
```

Again, a more detailed encoding for places is feasible if the `namesdates` module is included in the schema. The above examples assume that only the general purpose <name> element supplied in the core module is available.

## 16.3 Associating Contextual Information with a Text

This section discusses the association of the contextual information held in the header with the individual elements making up a TEI text or corpus. Contextual information is held in elements of various kinds within the TEI header, as discussed elsewhere in this section and in chapter *2. The TEI Header*. Here we consider what happens when different parts of a document need to be associated with different contextual information of the same type, for example when one part of a document uses a different encoding practice from another, or where one part relates to a different setting from another. In such situations, there will be more than one instance of a header element of the relevant type.

The TEI scheme allow for the following possibilities:

- A given element may appear in the corpus header only, in the header of one or more texts only, or in both places
- There may be multiple occurrences of certain elements in either the corpus or a text header.

To simplify the exposition, we deal with these two possibilities separately in what follows; however, they may be combined as desired.

### 16.3.1 Combining Corpus and Text Headers

A TEI-conformant document may have more than one header only in the case of a TEI corpus, which must have a header in its own right, as well as the obligatory header for each text. Every element specified in a corpus-header is understood as if it appeared within every text header in the corpus. An element specified in a text header but not in the corpus header supplements the specification for that text alone. If any element is specified in both corpus and text headers, the corpus header element is over-ridden for that text alone.

The <titleStmt> for a corpus text is understood to be prefixed by the <titleStmt> given in the corpus header. All other optional elements of the <fileDesc> should be omitted from an individual corpus text header unless they differ from those specified in the corpus header. All other header elements behave identically, in the manner documented below. This facility makes it possible to state once for all in the corpus header each piece of contextual information which is common to the whole of the corpus, while still allowing for individual texts to vary from this common denominator.

For example, the following schematic shows the structure of a corpus comprising three texts, the first and last of which share the same encoding description. The second one has its own encoding description.

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
<!-- corpus file description-->
    </fileDesc>
    <encodingDesc>
<!-- default encoding description -->
    </encodingDesc>
    <revisionDesc>
<!-- corpus revision description -->
    </revisionDesc>
  </teiHeader>
  <TEI xmlns="http://www.tei-c.org/ns/1.0">
    <teiHeader>
      <fileDesc>
<!-- file description for this corpus text -->
      </fileDesc>
    </teiHeader>
    <text>
<!-- first corpus text -->
    </text>
  </TEI>
  <TEI xmlns="http://www.tei-c.org/ns/1.0">
    <teiHeader>
      <fileDesc>
```

```
<!-- file description for this corpus text -->
     </fileDesc>
     <encodingDesc>
<!-- encoding description for this corpus
text, over-riding the default -->
     </encodingDesc>
   </teiHeader>
   <text>
<!-- second corpus text -->
   </text>
  </TEI>
  <TEI xmlns="http://www.tei-c.org/ns/1.0">
    <teiHeader>
      <fileDesc>
<!-- file description for third corpus text -->
     </fileDesc>
   </teiHeader>
   <text>
<!-- third corpus text -->
   </text>
  </TEI>
</TEI>
```

### 16.3.2 Declarable Elements

Certain of the elements which can appear within a TEI header are known as *declarable elements*. These elements have in common the fact that they may be linked explicitly with a particular part of a text header or the corpus header by means of a *decls* attribute on that element. This linkage is used to over-ride the default association between declarations in the header and a corpus or corpus text. The only header elements which may be associated in this way are those which would not otherwise be meaningfully repeatable.

Declarable elements are all members of the class `att.declarable`; the corresponding declaring elements are all members of the class `att.declaring`.

`att.declarable` provides attributes for those elements in the TEI header which may be independently selected by means of the special purpose *decls* attribute.

- `@default` indicates whether or not this element is selected by default when its parent is selected.

`att.declaring` provides attributes for elements which may be independently associated with a particular declarable element within the header, thus overriding the inherited default for that element.

- `@decls` (declarations) identifies one or more *declarable elements* within the header, which are understood to apply to the element bearing this attribute and its content.

An alphabetically ordered list of declarable elements follows:

`<availability>` (availability) supplies information about the availability of a text, for example any restrictions on its use or distribution, its copyright status, any licence applying to it, etc.

`<bibl>` (bibliographic citation) contains a loosely-structured bibliographic citation of which the sub-components may or may not be explicitly tagged.

`<biblFull>` (fully-structured bibliographic citation) contains a fully-structured bibliographic citation, in which all components of the TEI file description are present.

`<biblStruct>` (structured bibliographic citation) contains a structured bibliographic citation, in which only bibliographic sub-elements appear and in a specified order.

`<broadcast>` (broadcast) describes a broadcast used as the source of a spoken text.

`<correction>` (correction principles) states how and under what circumstances corrections have been made in the text.

`<correspDesc>` (correspondence description) contains a description of the actions related to one act of correspondence.

`<editorialDecl>` (editorial practice declaration) provides details of editorial principles and practices applied during the encoding of a text.

`<equipment>` (equipment) provides technical details of the equipment and media used for an audio or video recording used as the source for a spoken text.

`<geoDecl>` (geographic coordinates declaration) documents the notation and the datum used for geographic coordinates expressed as content of the <geo> element elsewhere within the document.

`<hyphenation>` (hyphenation) summarizes the way in which hyphenation in a source text has been treated in an encoded version of it.

`<interpretation>` (interpretation) describes the scope of any analytic or interpretive information added to the text in addition to the transcription.

`<langUsage>` (language usage) describes the languages, sublanguages, registers, dialects, etc. represented within a text.

`<listApp>` (list of apparatus entries) contains a list of apparatus entries.

`<listBibl>` (citation list) contains a list of bibliographic citations of any kind.

`<listEvent>` (list of events) contains a list of descriptions, each of which provides information about an identifiable event.

`<listNym>` (list of canonical names) contains a list of nyms, that is, standardized names for any thing.

`<listObject>` (list of objects) contains a list of descriptions, each of which provides information about an identifiable physical object.

`<listOrg>` (list of organizations) contains a list of elements, each of which provides information about an identifiable organization.

`<listPerson>` (list of persons) contains a list of descriptions, each of which provides information about an identifiable person or a group of people, for example the participants in a language interaction, or the people referred to in a historical source.

`<listPlace>` (list of places) contains a list of places, optionally followed by a list of relationships (other than containment) defined amongst them.

`<metDecl>` (metrical notation declaration) documents the notation employed to represent a metrical pattern when this is specified as the value of a *met*, *real*, or *rhyme* attribute on any structural element of a metrical text (e.g. <lg>, <l>, or <seg>).

`<normalization>` (normalization) indicates the extent of normalization or regularization of the original source carried out in converting it to electronic form.

`<particDesc>` (participation description) describes the identifiable speakers, voices, or other participants in any kind of text or other persons named or otherwise referred to in a text, edition, or metadata.

`<projectDesc>` (project description) describes in detail the aim or purpose for which an electronic file was encoded, together with any other relevant information concerning the process by which it was assembled or collected.

`<punctuation>` specifies editorial practice adopted with respect to punctuation marks in the original.

`<quotation>` (quotation) specifies editorial practice adopted with respect to quotation marks in the original.

`<recording>` (recording event) provides details of an audio or video recording event used as the source of a spoken text, either directly or from a public broadcast.

`<refsDecl>` (references declaration) specifies how canonical references are constructed for this text.

`<samplingDecl>` (sampling declaration) contains a prose description of the rationale and methods used in selecting texts, or parts of a text, for inclusion in the resource.

`<scriptStmt>` (script statement) contains a citation giving details of the script used for a spoken text.

`<segmentation>` (segmentation) describes the principles according to which the text has been segmented, for example into sentences, tone-units, graphemic strata, etc.

`<seriesStmt>` (series statement) groups information about the series, if any, to which a publication belongs.

`<settingDesc>` (setting description) describes the setting or settings within which a language interaction takes place, or other places otherwise referred to in a text, edition, or metadata.

`<sourceDesc>` (source description) describes the source(s) from which an electronic text was derived or generated, typically a bibliographic description in the case of a digitized text, or a phrase such as 'born digital' for a text which has no previous existence.

`<stdVals>` (standard values) specifies the format used when standardized date or number values are supplied.

`<styleDefDecl>` (style definition language declaration) specifies the name of the formal language in which style or renditional information is supplied elsewhere in the document. The specific version of the scheme may also be supplied.

`<textClass>` (text classification) groups information which describes the nature or topic of a text in terms of a standard classification scheme, thesaurus, etc.

`<textDesc>` (text description) provides a description of a text in terms of its situational parameters.

`<xenoData>` (non-TEI metadata) provides a container element into which metadata in non-TEI formats may be placed.

Each of the above elements is repeatable within a single header; that is, there may be more than one instance of any declarable element type at a given level. When this occurs, the following rules apply. For each different type of declarable element which occurs more than once:

- each must bear a unique identifier
- when occurring within the same parent element, exactly one element must be specified as the default, by having a *default* attribute with the value "true".

In the following example, an editorial declaration contains two possible <correction> policies, one identified as CorPol1 and the other as CorPol2. Since there are two, one of them (in this case CorPol1) should be specified as the default:

```
<editorialDecl>
  <correction  xml:id="CorPol1"
    default="true">
    <p> ... </p>
  </correction>
  <correction  xml:id="CorPol2">
    <p> ... </p>
  </correction>
  <normalization  xml:id="n1">
    <p> ... </p>
    <p> ... </p>
  </normalization>
</editorialDecl>
```

For texts associated with the header in which this declaration appears, correction method CorPol1 will be assumed, unless they explicitly state otherwise. Here is the structure of a text in which a division states otherwise:

```
<text>
  <body>
    <div1  n="D1"/>
    <div1  n="D2"  decls="#CorPol2"/>
    <div1  n="D3"/>
  </body>
</text>
```

In this case, the contents of the divisions D1 and D3 will both use correction policy CorPol1, and those of division D2 will use correction policy CorPol2.

The *decls* attribute is defined for any element which is a member of the class *declaring*. This includes the major structural elements <text>, <group>, and , as well as smaller structural units, down to the level of paragraphs

in prose, individual utterances in spoken texts, and entries in dictionaries. However, TEI recommended practice is to limit the number of multiple declarable elements used by a document as far as possible, for simplicity and ease of processing.

The identifier or identifiers to which the *decls* attribute points must follow two further restrictions:

- When *decl* references an element that itself contains multiple elements of the same type, only the children elements with *default* set to "true" are considered referenced.
- Each element specified, explicitly or implicitly, by the list of identifiers must be of a different kind.

To demonstrate how these rules operate, we now expand our earlier example slightly:

```
<encodingDesc>
  <editorialDecl xml:id="ED1" default="true">
    <correction xml:id="C1A" default="true">
      <p> ... </p>
    </correction>
    <correction xml:id="C1B">
      <p> ... </p>
    </correction>
    <normalization xml:id="N1">
      <p> ... </p>
      <p> ... </p>
    </normalization>
  </editorialDecl>
  <editorialDecl xml:id="ED2">
    <correction xml:id="C2A" default="true">
      <p> ... </p>
    </correction>
    <correction xml:id="C2B">
      <p> ... </p>
    </correction>
    <normalization xml:id="N2A">
      <p> ... </p>
    </normalization>
    <normalization xml:id="N2B"
      default="true">
      <p> ... </p>
    </normalization>
  </editorialDecl>
</encodingDesc>
```

This encoding description now has two editorial declarations, identified as ED1 (the default) and ED2. For texts not specifying otherwise, ED1 will apply. If ED1 applies, correction method C1A and normalization method N1 apply, since these are the specified defaults within ED1. In the same way, for a text specifying *decls* as '#ED2', correction C2A, and normalization N2B will apply.

A finer grained approach is also possible. A text might specify <text decls='#C2B #N2A'>, to 'mix and match' declarations as required. A tag such as <text decls='#ED1 #ED2'> would (obviously) be illegal, since it includes two elements of the same type; a tag such as <text decls='#ED2 #C1A'> is also illegal, since in this context #ED2 is synonymous with the defaults for that editorial declaration, namely #C2A #N2B, resulting in a list that identifies two correction elements (C1A and C2A).

### 16.3.3 Summary

The rules determining which of the declarable elements are applicable at any point may be summarized as follows:

1. If there is a single occurrence of a given declarable element in a corpus header, then it applies by default to all elements within the corpus.
2. If there is a single occurrence of a given declarable element in the text header, then it applies by default to all elements of that text irrespective of the contents of the corpus header.

3. Where there are multiple occurrences of declarable elements within a text's header or its corpus header,
   - each must have a unique value specified as the value of its *xml:id* attribute;
   - one and only one must bear a *default* attribute with the value true (or 1).
4. It is a semantic error for an element to be associated with more than one occurrence of any declarable element.
5. Selecting an element which contains multiple occurrences of a given declarable element is semantically equivalent to selecting only those contained elements which are specified as defaults.
6. An association made by one element applies by default to all of its descendants.

## 16.4 Linguistic Annotation of Corpora

Language corpora often include analytic encodings or annotations, designed to support a variety of different views of language. The present Guidelines do not advocate any particular approach to linguistic annotation (or 'tagging'); instead a number of general analytic facilities are provided which support the representation of most forms of annotation in a standard and self-documenting manner. Analytic annotation is of importance in many fields, not only in corpus linguistics, and is therefore discussed in general terms elsewhere in the Guidelines.[94] The present section presents informally some particular applications of these general mechanisms to the specific practice of corpus linguistics.

### 16.4.1 Levels of Analysis

By *linguistic annotation* we mean here any annotation determined by an analysis of linguistic features of the text, excluding as borderline cases both the formal structural properties of the text (e.g. its division into chapters or paragraphs) and descriptive information about its context (the circumstances of its production, its genre, or medium). The structural properties of any TEI-conformant text should be represented using the structural elements discussed elsewhere in these Guidelines, for example in chapters *3. Elements Available in All TEI Documents* and *4. Default Text Structure*. The contextual properties of a TEI text are fully documented in the TEI header, which is discussed in chapter *2. The TEI Header*, and in section *16.2. Contextual Information* of the present chapter.

Other forms of linguistic annotation may be applied at a number of levels in a text. A code (such as a word-class or part-of-speech code) may be associated with each word or token, or with groups of such tokens, which may be continuous, discontinuous, or nested. A code may also be associated with relationships (such as cohesion) perceived as existing between distinct parts of a text. The codes themselves may stand for discrete non-decomposable categories, or they may represent highly articulated bundles of textual features. Their function may be to place the annotated part of the text somewhere within a narrowly linguistic or discoursal domain of analysis, or within a more general semantic field, or any combination drawn from these and other domains.

The manner by which such annotations are generated and attached to the text may be entirely automatic, entirely manual, or a mixture. The ease and accuracy with which analysis may be automated may vary with the level at which the annotation is attached. The method employed should be documented in the <interpretation> element within the encoding description of the TEI header, as described in section *2.3.3. The Editorial Practices Declaration*. Where different parts of a corpus have used different annotation methods, the *decls* attribute should be used to indicate the fact, as further discussed in section *16.3. Associating Contextual Information with a Text*.

An extended example of one form of linguistic analysis commonly practised in corpus linguistics is given in section *18.4. Linguistic Annotation*.

## 16.5 Recommendations for the Encoding of Large Corpora

These Guidelines include proposals for the identification and encoding of a far greater variety of textual features and characteristics than is likely to be either feasible or desirable in any one language corpus, however large and ambitious. The reasoning behind this universal approach is further discussed in chapter *iv About These Guidelines*. For most large-scale corpus projects, it will therefore be necessary to determine a subset of TEI recommended elements appropriate to the anticipated needs of the project, as further discussed in chapter *24.3. Customization*; these mechanisms include the ability to exclude selected element types, add new element types, and change the names of existing elements. A discussion of the implications of such changes for TEI conformance is provided in chapter *24.4. Conformance*.

[94] See in particular chapters *17. Linking, Segmentation, and Alignment*, *18. Simple Analytic Mechanisms*, and *19. Feature Structures*.

Because of the high cost of identifying and encoding many textual features, and the difficulty in ensuring consistent practice across very large corpora, encoders may find it convenient to divide the set of elements to be encoded into the following four categories:

**required** texts included within the corpus will always encode textual features in this category, should they exist in the text

**recommended** textual features in this category will be encoded wherever economically and practically feasible; where present but not encoded, a note in the header should be made.

**optional** textual features in this category may or may not be encoded; no conclusion about the absence of such features can be inferred from the absence of the corresponding element in a given text.

**proscribed** textual features in this category are deliberately not encoded; they may be transcribed as unmarked up text, or represented as <gap> elements, or silently omitted, as appropriate.

## 16.6 Module for Language Corpora

The module described in this chapter makes available the following components:

**Module corpus: Corpus texts**

- *Elements defined*: activity channel constitution derivation domain factuality interaction locale particDesc preparedness purpose setting settingDesc textDesc

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 17 Linking, Segmentation, and Alignment

This chapter discusses a number of ways in which encoders may represent analyses of the structure of a text which are not necessarily linear or hierarchic. The module defined by this chapter provides for the following common requirements:

- to link disparate elements using the *xml:id* attribute (section *17.1. Links*);
- to link disparate elements without using the *xml:id* attribute (sections *17.2.1. Pointing Elsewhere* and *17.2.4. TEI XPointer Schemes*);
- to segment text into elements convenient for the encoder and to mark arbitrary points within documents (section *17.3. Blocks, Segments, and Anchors*);
- to represent correspondence or alignment among groups of text elements, both those with content and those which are empty (section *17.5. Correspondence and Alignment*);[95]
- to synchronize elements of a text, that is to represent temporal correspondences and alignments among text elements (section *17.4. Synchronization*) and also to align them with specific points in time (section *17.4.2. Placing Synchronous Events in Time*);
- to specify that one text element is identical to or a copy of another (section *17.6. Identical Elements and Virtual Copies*);
- to aggregate possibly noncontiguous elements (section *17.7. Aggregation*);
- to specify that different elements are alternatives to one another and to express preferences among the alternatives (section *17.8. Alternation*);
- to store markup separately from the data it describes or is related to (section *17.9. Stand-off Markup*);
- to associate segments of a text with interpretations or analyses of their significance (section *17.12. Connecting Analytic and Textual Markup*);
- to group together elements used to provide stand-off annotation, including contextual information (section *17.10. The standOff Container*).

These facilities all use the same set of techniques based on the W3C XPointer framework (?) This provides a variety of *schemes*; the most convenient of which, and that recommended by these Guidelines, makes use of the global *xml:id* attribute, as defined in section *1.3.1.1. Global Attributes*, and introduced in the section of *v A Gentle Introduction to XML* titled *v.6.2 Identifiers and Indicators*. When the `linking` module is included in a schema, the attribute class `att.global` is extended to include eight additional attributes to support the various kinds of linking listed above. Each of these attributes is introduced in the appropriate section below. In addition, for many of the topics discussed, a choice of methods of encoding is offered, ranging from simple but less general ones, which use attribute values only, to more elaborate and more general ones, which use specialized elements.

## 17.1 Links

We say that one element *points* to others if the first has an attribute whose value is a reference to the others: such an element is called a *pointer element*, or simply a *pointer*. Among the pointers that have been introduced up to this point in these Guidelines are <note>, <ref>, and <ptr>. These elements all indicate an association between one place in the document (the location of the pointer itself) and one or more others (the elements whose identifiers are specified by the pointer's *target* attribute). The module described in this chapter introduces a variation on this basic kind of pointer, known as a *link*, which specifies both 'ends' of an association. In addition, we define a syntax for representing locations in a document by a variety of means not dependent on the use of *xml:id* attributes.

### 17.1.1 Pointers and Links

In section *3.7. Simple Links and Cross-References* we introduced the simplest pointer elements, <ptr> and <ref>. Here we introduce additionally the <link> element, which represents an association between two (or more) locations by specifying each location explicitly. Its own location is irrelevant to the intended linkage. All three

[95] We use the term *alignment* as a special case for the more general notion of correspondence. Using A as a short form for 'an element with its attribute *xml:id* set to the value A', and suppose elements A1, A2, and A3 occur in that order and form one group, while elements B1, B2, and B3 occur in that order and form another group. Then a relation in which A1 corresponds to B1, A2 corresponds to B2, and A3 corresponds to B3 is an alignment. On the other hand, a relation in which A1 corresponds to B2, B1 to C2, and C1 to A2 is not an alignment.

elements use the attribute *target*, provided by the `att.pointing` class as a means of indicating the location or locations referenced or pointed to.

`att.pointing` provides a set of attributes used by all elements which point to other elements by means of one or more URI references.

`@target` specifies the destination of the reference by supplying one or more URI References.

`<link>` (link) defines an association or hypertextual link among elements or passages, of some type not more precisely specifiable by other elements.

The <ptr> element may be called a 'pure pointer', because its primary function is simply to point. A pointer sets up a *connection* between an element (which, in the case of a pure pointer, is simply a location in a document), and one or more others, known collectively as its *target*. The <ptr> and <ref> elements point, conceptually, at a single target, even if that target may be discontinuous in the document. The <link> element specifies at least two targets and represents an association between them, independent of its own location.

These three elements also share a common set of attributes, derived from the `att.pointing` and `att.typed` classes:

`att.pointing` provides a set of attributes used by all elements which point to other elements by means of one or more URI references.

`@evaluate` (evaluate) specifies the intended meaning when the target of a pointer is itself a pointer.

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.

`@subtype` (subtype) provides a sub-categorization of the element, if needed.

Double connection among elements could also be expressed by a combination of pointer elements, for example, two <ptr> elements, or one <ptr> element and one <note> element. All that is required is that the value of the *target* (or other pointing) attribute of the one be the value of the *xml:id* attribute of the other. What the <link> element accomplishes is the handling of double connection by means of a single element. Thus, in the following encoding:

```
<ptr  xml:id="sa-p1"  target="#sa-p2"/>
<ptr  xml:id="sa-p2"  target="#sa-p1"/>
```

sa-p1 points to sa-p2, and sa-p2 points to sa-p1. This is logically equivalent to the more compact encoding:

```
<link  target="#sa-p1 #sa-p2"/>
```

As noted elsewhere, the *target* attribute may take as value one or more URI reference. In the simplest case, each such reference will indicate an element in the current document (or in some other document), for example by supplying the value used for its global *xml:id* attribute. It may however carry as value any form of URI, such as a URL pointing to some other document or location on the Internet. Pointing or linking to external documents and pointing and linking where identifiers are not available is described below in section *17.2. Pointing Mechanisms*.

### 17.1.2 Using Pointers and Links

As an example of the use of mechanisms which establish connections among elements, consider the practice (common in 18th century English verse and elsewhere) of providing footnotes citing parallel passages from classical authors. Such footnotes can of course simply be encoded using the <note> element (see section *3.9. Notes, Annotation, and Indexing*) without a *target* attribute, placed adjacent to the passage to which the note refers:[96]

[96] The *type* attribute on the note is used to classify the notes using the typology established in the Advertisement to the work: 'The *Imitations* of the Ancients are added, to gratify those who either never read, or may have forgotten them; together with some of the Parodies, and Allusions to the most excellent of the Moderns.' In the source text, the text of the poem shares the page with two sets of notes, one headed 'Remarks' and the other 'Imitations'.

Their full-fed Heroes, their pacific May'rs,
Their annual trophies, and their monthly wars.
Tho' long my Party built on me their hopes,
For writing pamphlets, and for burning Popes;
(Diff'rent our parties, but with equal grace
The Goddeſs ſmiles on Whig and Tory race,
'Tis the ſame rope at ſev'ral ends they twiſt,
To Dulneſs, Ridpath is as dear as Miſt.)
Yet lo! in me what authors have to brag on!
Reduc'd at laſt to hiſs in my own dragon.
Avert it, Heav'n! that thou or Cibber e'er
Should wag two ſerpent tails in Smithfield fair.
Like the vile ſtraw that's blown about the ſtreets
The needy Poet ſticks to all he meets,
Coach'd, carted, trod upon, now looſe, now faſt,
In the Dog's tail his progreſs ends at laſt.

*REMARKS.*

VERSE 280. Annual trophies, on the *Lord Mayor's Day*; and monthly wars, in the *Artillery Ground*.

VERSE 281. *Tho' long my Party.*] *Settle*, like moſt Party-writers, was very uncertain in his political principles. He was employ'd to hold the pen in the *Character* of a *Popiſh ſucceſſor*, but afterwards printed his *Narrative* on the contrary ſide. He had managed the Ceremony of a famous Pope-burning on *Nov.* 17, 1680: then became a Trooper of King *James's* army at *Hounſlow-heath*: After the Revolution he kept a Booth at *Bartlemew-fair*, where in his Droll call'd St. *George for England*, he acted in his old age in a Dragon of green leather of his own invention. He was at laſt taken into the Charter-houſe, and there dyed, aged about 60 years.

V. 285. *To Dulneſs*, Ridpath *is as dear as* Miſt.] *George Ridpath*, author for ſeveral years of the *Flying-Poſt*, a Whig-paper; *Nathaniel Miſt*, publiſher of the Weekly Journal, a Tory-paper.

*IMITATIONS.*

VERSE 283-84. —— *With equal grace*
*Our Goddeſs ſmiles on Whig and Tory race.*]
Virg. Æn. 10.

*Tros Rutulusve fuat, nullo diſcrimine habebo.*
—— *Rex Jupiter omnibus idem.*

P

```
<l>(Diff'rent our parties, but with equal grace</l>
<l>The Goddess smiles on Whig and Tory race,</l>
<l>
  <note  type="imitation"  place="bottom">
    <bibl>Virg. Æn. 10.</bibl>
    <quote>
      <l>Tros Rutulusve fuat; nullo discrimine habebo.</l>
      <l>—— Rex Jupiter omnibus idem.</l>
    </quote>
  </note>'Tis the same rope at sev'ral ends they twist,
</l>
<l>To Dulness, Ridpath is as dear as Mist)</l>
```

This use of the <note> element can be called *implicit pointing* (or *implicit linking*). It relies on the juxtaposition of the note to the text being commented on for the connection to be understood. If it is felt that the mere juxtaposition of the note to the text does not make it sufficiently clear exactly what text segment is being commented on (for example, is it the immediately preceding line, or the immediately preceding two lines, or what?), or if it is decided to place the note at some distance from the text, then the pointing or the linking must be made explicit. We now consider various methods for doing that.

Firstly, a <ptr> element might be placed at an appropriate point within the text to link it with the annotation:

```
<l>(Diff'rent our parties, but with equal grace</l>
<l>The Goddess smiles on Whig and Tory race,
<ptr  rend="unmarked"  target="#note3.284"/>
</l>
<l>'Tis the same rope at sev'ral ends they twist,</l>
<l>To Dulness, Ridpath is as dear as Mist)</l>
<note  xml:id="note3.284"  type="imitation"
  place="bottom">
  <bibl>Virg. Æn. 10.</bibl>
  <quote>
    <l>Tros Rutulusve fuat; nullo discrimine habebo.</l>
    <l>—— Rex Jupiter omnibus idem.</l>
  </quote>
</note>
```

The <note> element has been given an arbitrary identifier (note3.284) to enable it to be specified as the target of the pointer element. Because there is nothing in the text to signal the existence of the annotation, the *rend* attribute has been given the value unmarked.

Secondly, the *target* attribute of the <note> element can be used to point at its associated text, provided that an *xml:id* attribute has been supplied for the associated text:

```
<l  xml:id="L3.283">(Diff'rent our parties, but with equal grace</l>
<l  xml:id="L3.284">The Goddess smiles on Whig and Tory race,</l>
<l  xml:id="L3.285">'Tis the same rope at sev'ral ends they twist,</l>
<l  xml:id="L3.286">To Dulness, Ridpath is as dear as Mist)</l>
<!-- ... -->
```

Given this encoding of the text itself, we can now link the various notes to it. In this case, the note itself contains a pointer to the place in the text which it is annotating; this could be encoded using a <ref> element, which bears a *target* attribute of its own and contains a (slightly misquoted) extract from the text marked as a <quote> element:

```
<note type="imitation" place="bottom"
  target="#L3.284">
  <ref rend="sc" target="#L3.284">Verse 283–84.
  <quote>
     <l>——. With equal grace</l>
     <l>Our Goddess smiles on Whig and Tory race.</l>
   </quote>
  </ref>
  <bibl>Virg. Æn. 10.</bibl>
  <quote>
    <l>Tros Rutulusve fuat; nullo discrimine habebo.</l>
    <l>—— Rex Jupiter omnibus idem. </l>
  </quote>
</note>
```

Combining these two approaches gives us the following associations:

- a pointer within one line indicates the note
- the note indicates the line
- a pointer within the note indicates the line

Note that we do not have any way of pointing from the line itself to the note: the association is implied by containment of the pointer. We do not as yet have a true double link between text and note. To achieve that we will need to supply identifiers for the annotations as well as for the verse lines, and use a <link> element to associate the two. Note that the <ptr> element and the *target* attribute on the <note> may now be dispensed with:

```
<note xml:id="n3.284" type="imitation"
  place="bottom">
  <ref rend="sc" target="#L3.284">Verse 283–84.
  <quote>
     <l>——. With equal grace</l>
     <l>Our Goddess smiles on Whig and Tory race.</l>
   </quote>
  </ref>
  <bibl>Virg. Æn. 10.</bibl>
  <quote>
    <l>Tros Rutulusve fuat; nullo discrimine habebo.</l>
    <l>—— Rex Jupiter omnibus idem. </l>
  </quote>
</note>
<link target="#n3.284 #L3.284"/>
```

The *target* attribute of the <link> element here bears the identifier of the note followed by that of the verse line. We could also allocate an identifier to the reference within the note and encode the association between it and the verse line in the same way:

```
<note type="imitation" place="bottom">
  <ref rend="sc" xml:id="r3.284"
    target="#L3.284">Verse 283–84.
  <quote>
     <l>——. With equal grace</l>
     <l>Our Goddess smiles on Whig and Tory race.</l>
   </quote>
  </ref>
```

```
<!-- ... -->
</note>
<!-- ... -->
<link target="#r3.284 #L3.284"/>
```

Indeed, the two <link>s could be combined into one, as follows:

```
<link target="#n3.284 #r3.284 #L3.284"/>
```

### 17.1.3 Groups of Links

Clearly, there are many reasons for which an encoder might wish to represent a link or association between different elements. For some of them, specific elements are provided in these Guidelines; some of these are discussed elsewhere in the present chapter. The <link> element is a general purpose element which may be used for any kind of association. The element <linkGrp> may be used to group links of a particular type together in a single part of the document; such a collection may be used to represent what is sometimes referred to in the literature of Hypertext as a *web*, a term introduced by the Brown University FRESS project in 1969, and not to be confused with the World Wide Web.

`<linkGrp>` (link group) defines a collection of associations or hypertextual links.

As a member of the class `att.pointing.group`, this element shares the following attributes with other members of that class:

`att.pointing.group` provides a set of attributes common to all elements which enclose groups of pointer elements.

- `@domains` optionally specifies the identifiers of the elements within which all elements indicated by the contents of this element lie.
- `@targFunc` (target function) describes the function of each of the values of the *target* attribute of the enclosed <link>, <join>, or <alt> tags.

It is also a member of the `att.pointing` and `att.typed` classes, and therefore also carries the attributes specified in section *17.1.1. Pointers and Links* above, in particular the *type* attribute.

The <linkGrp> element provides a convenient way of establishing a default for the *type* attribute on a group of links of the same type: by default, the *type* attribute on a <link> element has the same value as that given for *type* on the enclosing <linkGrp>.

Typical software might hide a web entirely from the user, but use it as a source of information about links, which are displayed independently at their referenced locations. Alternatively, software might provide a direct view of the link collection, along with added functions for manipulating the collection, as by filtering, sorting, and so on. To continue our previous example, this text contains many other notes of a kind similar to the one shown above. Here are a few more of the lines to which annotations have to be attached, followed by the annotations themselves:

```
<l xml:id="L2.79">A place there is, betwixt earth, air and seas</l>
<l xml:id="L2.80">Where from Ambrosia, Jove retires for ease.</l>
<!-- ... -->
<l xml:id="L2.88">Sign'd with that Ichor which from Gods distills.</l>
<!-- ... -->
<note xml:id="n2.79" place="bottom"
  anchored="false">
  <bibl>Ovid Met. 12.</bibl>
  <quote xml:lang="la">
    <l>Orbe locus media est, inter terrasq; fretumq;</l>
    <l>Cœlestesq; plagas –</l>
  </quote>
</note>
<note xml:id="n2.88" place="bottom"
```

```
  anchored="false"> Alludes to <bibl>Homer, Iliad 5</bibl> ...
</note>
```

To avoid having to repeat the specification of *type* as imitation on each <note>, we may specify it once for all on a <linkGrp> element containing all links of this type.

```
<linkGrp  type="imitation">
  <link  target="#n2.79 #L2.79"/>
  <link  target="#n2.88 #L2.88"/>
  <link  target="#n3.284 #L3.284"/>
</linkGrp>
```

Additional information for applications that use <linkGrp> elements can be provided by means of special attributes. First, the *domains* attribute can be used to identify the text elements within which the individual targets of the links are to be found. Suppose that the text under discussion is organized into a <body> element, containing the text of the poem, and a <back> element containing the notes. Then the *domains* attribute can have as its value the identifiers of the <body> and the <back>, to enable an application to verify that the link targets are in fact contained by appropriate elements, or to limit its search space:

```
<!-- ... --><linkGrp  type="imitation"
  domains="#dunciad #dunnotes">
  <link  target="#n2.79 #L2.79"/>
  <link  target="#n2.88 #L2.88"/>
<!-- ... -->
  <link  target="#n3.284 #L3.284"/>
<!-- ... -->
</linkGrp>
```

Note that there must be a single parent element for each 'domain'; if some notes are contained by a section with identifier dunnotes, and others by a section with identifier dunimits, an intermediate pointer must be provided (as described in section *17.1.4. Intermediate Pointers*) within the <linkGrp> and its identifier used instead.

Next, the *targFunc* attribute can be used to provide further information about the role or function of the various targets specified for each link in the group. The value of the *targFunc* attribute is a list of names (formally, name tokens), one for each of the targets in the link; these names can be chosen freely by the encoder, but their significance should be documented in the encoding description in the header.[97] In the current example, we might think of the note as containing the *source* of the imitation and the verse line as containing the *goal* of the imitation. Accordingly, we can specify the <linkGrp> in the preceding example thus:

```
<linkGrp  type="imitation"
  domains="#dunciad #dunnotes"  targFunc="source goal">
  <link  target="#n2.79 #L2.79"/>
  <link  target="#n2.88 #L2.88"/>
<!-- ... -->
  <link  target="#n3.284 #L3.284"/>
<!-- ... -->
</linkGrp>
```

[97] Since no special element is provided for this purpose in the present version of these Guidelines, the information should be supplied as a series of paragraphs at the end of the <encodingDesc> element described in section *2.3. The Encoding Description.*

### 17.1.4 Intermediate Pointers

In the preceding examples, we have shown various ways of linking an annotation and a single verse line. However, the example cited in fact requires us to encode an association between the note and a *pair* of verse lines (lines 284 and 285); we call these two lines a *span*.

There are a number of possible ways of correcting this error: one could use the *target* attribute to indicate one end of the span and the special purpose *targetEnd* attribute on the <note> element to point to the other. Another possibility might be to create an element which represents the whole span itself and assign that an *xml:id* attribute, which can then be linked to the <note> and <ref> elements. This could be done using for example the <lg> element defined in section *3.13.1. Core Tags for Verse* or the 'virtual' <join> element discussed in section *17.7. Aggregation*.

A third possibility would be to use an 'intermediate pointer' as follows:

```
<ptr  xml:id="L3.283-284"
  target="#L3.283 #L3.284"/>
```

When the *target* attribute of a <ptr> or <ref> element specifies more than one element, the indicated elements are intended to be combined or aggregated in some way to produce the object of the pointer. (Such aggregation is however the task of a processing application, and cannot be defined simply by the markup). The *xml:id* attribute of the <ptr> then provides an identifier which can be linked to the <note> and <ref> elements:

```
<link  evaluate="all"
  target="#n3.284 #r3.284 #L3.283-284"/>
```

The all value of *evaluate* is used on the <link> element to specify that any pointer encountered as a target of that element is itself evaluated. If *evaluate* had the value none, the link target would be the pointer itself, rather than the objects it points to.

Where a <linkGrp> element is used to group a collection of <link> elements, any intermediate pointer elements used by those <link> elements should be included within the <linkGrp>.

## 17.2 Pointing Mechanisms

This section introduces more formally the pointing mechanisms available in the TEI. In addition to those discussed so far, the TEI provides methods of pointing:

- into documents other than the current document;
- to a particular element in a document other than the current document using its *xml:id*;
- to a particular element whether in the current document or not, using its position in the XML element tree;
- at arbitrary content in any XML document using TEI-defined XPointer schemes.

All TEI attributes used to point at something else are declared as having the datatype `teidata.pointer`, which is defined as a URI reference[98]; the cases so far discussed are all simple examples of a URI reference. Another familiar example is the mechanism used in XHTML to create represent hypertext links by means of the XHTML *href* attribute. A URI reference can reference the whole of an XML resource such as a document or an XML element, or a sub-portion of such a resource, identified by means of an appropriate *fragment identifier*. Technically speaking, the 'fragment identifier' is that portion of a URI reference following the first unescaped '#' character; in practice, it provides a means of accessing some part of the resource described by the URI which is less than the whole.

The first three of the following subsections provide only a brief overview and some examples of the W3C mechanisms recommended. More detailed information on the use of these mechanisms is readily available elsewhere.

### 17.2.1 Pointing Elsewhere

Like the ubiquitous if misnamed XHTML pointing attribute *href*, the TEI pointing attributes can point to a document that is not the current document (the one that contains the pointing element) whether it is in the same local filesystem as the current document, or on a different system entirely. In either case, the pointing can be accomplished absolutely

[98] The URI (Universal Resource Indicator) is defined in RFC 3986

(using the entire address of the target document) or relatively (using an address relative to the current base URI in force). The 'current base URI' is defined according to Marsh and Tobin 2009. If there is none, the base URI is that of the current document. In common practice the current base URI in force is likely to be the value of the *xml:base* attribute of the closest ancestor that has one. However this may not be the case, since *xml:base* attributes are accumulated through the hierarchy by concatenation of path segments, beginning at the top of the hierarchy and proceeding down to the context node.

The following example demonstrates an absolute URI reference that points to a remote document:

```
The current base URI in force is as defined in the
W3C <ref  target="https://www.w3.org/TR/xmlbase/">XML
Base</ref> recommendation.
```

This example points explicitly to a location on the Web, accessible via HTTP. Suppose however that we wish to access a document stored locally in a file. Again we will supply an absolute URI reference, but this time using a different protocol:

```
This Debian package is distributed under the terms
of the <ref  target="file:///usr/share/common-licenses/GPL-2">GNU General Public
License</ref>.
```

In the following example, we use a relative URI reference to point to a local document:

```
<figure  rend="float fullpage">
  <graphic  url="Images/compic.png"/>
  <figDesc>The figure shows the page from the <title>Orbis
        pictus</title> of Comenius which is discussed in the text.</figDesc>
</figure>
```

Since no *xml:base* is specified here, the location of the resource `Images/compic.png` is determined relative to the resource indicated by the current base URI, which is the current document.

In the following example, however, we first change the current base URI by setting a new value for *xml:base*. The resource required is then identified by means of a relative URI:

```
<div  type="chap"
  xml:base="http://classics.mit.edu/">
  <head>On Ancient Persian Manners</head>
  <p>In the very first story of <ref  target="Sadi/gulistan.2.i.html">
     <title>The Gulistan of
           Sa'di</title>
   </ref>,
    Sa'di relates moral advice worthy of Miss Minners ...</p>
<!-- ... -->
</div>
```

As noted above, the current base URI is found on the nearest ancestor. It is technically possible to use *xml:base* as a means to shorten URIs, but this usage is not recommended. Abbreviated pointers provide a more flexible and consistent method for creating shorthand links.

### 17.2.2 Pointing Locally

A *shorthand pointer*, in which the URI consists only of **#** followed by the value of an *xml:id* acts as a pointer to the element in the current document with that *xml:id*, as in the following example.

```
<div  type="section"  xml:id="sect106">
<!-- ... -->
</div>
<div  type="section"  n="107"  xml:id="sect107">
  <head>Limitations on exclusive rights: Fair use</head>
  <p>Notwithstanding the provisions of
  <ref  target="#sect106">section 106</ref>, the fair use of a
     copyrighted work, including such use by reproduction in copies
     or phonorecords or by any other means specified by that section,
     for purposes such as criticism, comment, news reporting,
     teaching (including multiple copies for classroom use),
     scholarship, or research, is not an infringement of copyright.
     In determining whether the use made of a work in any particular
     case is a fair use the factors to be considered shall
     include –
  <list  rend="bulleted">
      <item  n="(1)">the purpose and character of the use, including
            whether such use is of a commercial nature or is for nonprofit
            educational purposes;</item>
      <item  n="(2)">the nature of the copyrighted work;</item>
      <item  n="(3)">the amount and substantiality of the portion
            used in relation to the copyrighted work as a whole;
            and</item>
      <item  n="(4)">the effect of the use upon the potential market
            for or value of the copyrighted work.</item>
    </list>
     The fact that a work is unpublished shall not itself bar a
     finding of fair use if such finding is made upon consideration
     of all the above factors.</p>
</div>
```

This method of pointing, by referring to the *xml:id* of the target element as a bare name only (e.g., #sect106) is the simplest and often the best approach where it can be applied, i.e. where both the source element and target element are in the same XML document, and where the target element carries an identifier. It is the method used extensively in previous sections of this chapter and elsewhere in these Guidelines.

### 17.2.3 Using Abbreviated Pointers

Even in the case of relative links on the local file system, *ref* or *target* attributes may become quite lengthy and make XML code difficult to read. To deal with this problem, the TEI provides a useful method of using abbreviated pointers and documenting a way to dereference them automatically.

Imagine a project which has a large collection of XML documents organized like this:

- anthology
  - poetry
    - `poem.xml`
  - prose
    - `novel.xml`
- references
  - people
    - `personography.xml`

If you want to link a <name> in the `novel.xml` file to a <person> in the `personography.xml` file, the link will look like this:

```
<name ref="../../references/people/personography.xml#fred">Fred</name>
```

If there are many names to tag in a single paragraph, the XML encoding will be congested, and such lengthy links are prone to typographical error. In addition, if the project organization is changed, every relative link will have to be found and altered.

One way to deal with this is to use what is often referred to as a ”magic token”. You could make such links using the *key* attribute:

```
<name key="fred">Fred</name>
```

and document the meaning of the key using (for instance) a <taxonomy> element in the TEI header, as described in *3.6.1. Referring Strings*. However, such a link cannot be mechanically processed by an external system that does not know how to interpret it; a human will have to read the header explanation and write code explicitly to reconstruct the intended link.

A more robust alternative is to use a *private URI scheme*. This is a method of constructing a simple, key-like token which functions as a `teidata.pointer`, and can therefore be used as the value of any attribute which has that datatype, such as *ref* and *target*. Such a scheme consists of a prefix with a colon, and then a value. You might, for example, use the prefix psn (for ”person”), and structure your name tags like this:

```
<name ref="psn:fred">Fred</name>
```

How is this different from a ‘magic token’? Essentially, it isn’t, except that TEI provides a structured method of dereferencing it (turning it into a computable path, such as ../../references/people/personography.xml#fred) by means of a declaration inside <encodingDesc> in the TEI header, using the elements and attributes for prefix declaration:

`<listPrefixDef>` (list of prefix definitions) contains a list of definitions of prefixing schemes used in `teidata.pointer` values, showing how abbreviated URIs using each scheme may be expanded into full URIs.

`<prefixDef>` (prefix definition) defines a prefixing scheme used in `teidata.pointer` values, showing how abbreviated URIs using the scheme may be expanded into full URIs.

- `@ident` supplies a name which functions as the prefix for an abbreviated pointing scheme such as a private URI scheme. The prefix constitutes the text preceding the first colon.

`att.patternReplacement` provides attributes for regular-expression matching and replacement.

- `@matchPattern` specifies a regular expression against which the values of other attributes can be matched.
- `@replacementPattern` specifies a ‘replacement pattern’, that is, the skeleton of a relative or absolute URI containing references to groups in the *matchPattern* which, once subpattern substitution has been performed, complete the URI.

This is how you might document a private URI scheme using the psn: prefix:

```
<listPrefixDef>
  <prefixDef ident="psn"
    matchPattern="([a-z]+)"
    replacementPattern="../../references/people/personography.xml#$1">
    <p> In the context of this project, private URIs with the prefix
        "psn" point to <gi>person</gi> elements in the project's
        personography.xml file.
    </p>
  </prefixDef>
</listPrefixDef>
```

This specifies that where a `teidata.pointer` value is constructed with a psn: prefix, a regular-expression replace operation can be performed on it to construct the full or relative URI to the target document or fragment. <listPrefixDef> is a child of <encodingDesc>, and it contains any number of <prefixDef> elements. Each <prefixDef> element provides a method of dereferencing or expanding an abbreviated pointer, based on a regular expression. The *ident* attribute specifies the prefix to which the expansion applies (without the colon). The *matchPattern* attribute contains a regular expression which is matched against the component of the pointer following the first colon, and the *replacementPattern* provides the string which will be used as a replacement. In this example, using psn:fred, the value fred would be matched by the *matchPattern*, and also captured (through the parentheses in the regular expression); it would then be replaced by the value ../../references/people/personography.xml#fred (with the the $1 in the *replacementPattern* being replaced by the captured value). The <p> element inside the <prefixDef> can be used to provide a human-readable explanation of the usage of this prefix.

Through this mechanism, any processor which encounters a `teidata.pointer` with a protocol unknown to it can check the <listPrefixDef> in the header to see if there is an available expansion for it, and if there is, it can automatically provide the expansion and generate a full or relative URI.

For any given prefix, it may be useful to supply more than one expansion. For instance, in addition to pointing at the <person> element in the personography file, it might also be useful to point to an external source which is available on the network, representing the same information in a different way. So there might be a second <prefixDef> like this:

```
<prefixDef  ident="psn"
  matchPattern="([a-z]+)"
  replacementPattern="http://www.example.com/personography.html#$1">
  <p> Private URIs with the prefix "psn" can be converted to point
     to a fragment on the Personography page of the project Website.
  </p>
</prefixDef>
```

Any number of <prefixDef> elements may be provided for the same prefix. A processor may decide to process one or all of them; if it processes only one, it should choose the first one with the correct *ident* value, so the primary or most important <prefixDef> for any given prefix should appear first in its parent <listPrefixDef>.

When creating private URI schemes, it is recommended that you avoid using any existing registered prefix. A list of registered prefixes is maintained by IANA at `https://www.iana.org/assignments/uri-schemes.html`.

Note that this mechanism can also be used to dereference other abbreviated pointing systems which are based on prefixes, such as Tag URIs.

The *matchPattern* and *replacementPattern* attributes are also used in dereferencing canonical reference patterns, and further examples of the use of regular expressions are shown in *17.2.5. Canonical References*.

## 17.2.4 TEI XPointer Schemes

The pointing schemes described in this chapter are part of a number of such schemes envisaged by the W3C, which together constitute a framework for addressing data within XML documents, known as the XPointer Framework (Grosso et al 2003). This framework permits the definition of many other named addressing methods, each of which is known as an *XPointer Scheme*. The W3C has predefined a set of such schemes, and maintains a register for their expansion.

One important scheme, also defined by the W3C, and recommended by these Guidelines is the xpath() pointer scheme, which allows for any part of an XML structure to be selected using the syntax defined by the XPath specification. This is further discussed below, *17.2.4.2. xpath()*. These Guidelines also define six other pointer schemes, which provide access to parts of an XML document such as points within data content or stretches of data content. These additional TEI pointer schemes are defined in sections *17.2.4.3. left()* to *17.2.4.8. match()* below.

### 17.2.4.1 Introduction to TEI Pointers

Before discussing the TEI pointer schemes, we introduce slightly more formally the terminology used to define them. So far, we have discussed only ways of pointing at components of the XML information set node such as elements and attributes. However, there is often a need in text analysis to address additional types of location such as the

'point' locations *between* 'nodes', and 'sequences' that may arbitrarily cross the boundaries of nodes in a document. The content of an XML document is organized sequentially as well as hierarchically, and it makes sense to consider ranges of characters within a document independently of the nodes to which they belong. From the perspective of most of the pointer schemes discussed below, a TEI document is a tree structure superimposed upon a character stream. Nodes are entities available only in the tree, while points are available only in the stream. For this reason, the schemes below that rely upon character positions (`string-index()`, `string-range()`, and `match()`) cannot take nodes into account. Conversely, XPath (disregarding functions that return atomic values) is a method for locating nodes in the tree and treats those nodes as indivisible units, meaning it is unable to address parts of nodes in their document context.

The TEI pointer scheme thus distinguishes the following kinds of object:

**Node** A node is an instance of one of the node kinds defined in the XQuery and XPath Data Model 3.1. It represents a single item in the XML information set for a document. For pointing purposes, the only nodes that are of interest are Text Nodes, Element Nodes, and Attribute Nodes.

**Sequence** A Sequence follows the definition in the XPath 3.1 Data Model, with one alteration. A Sequence is an ordered collection of zero or more items, where an item is either a node or a partial text node.

**Text Stream** A Text Stream is the concatenation of the text nodes in a document and behaves as though all tags had been removed. A text stream begins at a reference node and encompasses all of the text inside that node (if any) and all the text following it in document order. In XPath terms, this would encompass all of the text nodes beginning at a particular node, and following it on the following axis.

**Point** A Point represents a dimensionless point between nodes or characters in a document. Every point is adjacent to either characters or elements, and never to another point. Points can only be referenced in relation to an element or text node in the document (i.e. something addressable by either an XPath or a fragment identifier). Points occur either immediately before or after an element, or at a numbered position inside a text stream. Position zero in the stream would be immediately before the first character. Note that points within attribute values cannot mark the beginning or end of a range extending beyond the attribute value, because points indicate a position within a document. Since attribute nodes are by definition un-ordered, they cannot be said to have a fixed position.

The TEI recommends the following seven pointer schemes:

**xpath()** Addresses a node or node sequence using the XPath syntax. (*17.2.4.2. xpath()*)

**left() and right()** addresses the point before (left) or after (right) a node or node sequence (*17.2.4.3. left()* and *17.2.4.4. right()*)

**string-index()** addresses a point inside a text node (*17.2.4.5. string-index()*

**range()** addresses the range between two points (*17.2.4.6. range()*)

**string-range()** addresses a range of a specified length starting from a specified point (*17.2.4.7. string-range()*)

**match()** addresses a range which matches a specified string within a node (*17.2.4.8. match()*)

The xpath() scheme refers to the existing XPath specification which is adopted with one modification: the default namespace for any XPath used as a parameter to this scheme is assumed to be the TEI namespace `http://www.tei-c.org/ns/1.0`.

The other six schemes overlap in functionality with a W3C draft specification known as the XPointer scheme draft, but are individually much simpler. At the time of this writing, there is no current or scheduled activity at the W3C towards revising this draft or issuing it as a recommendation.

**A note on namespaces**: The W3C defines an xmlns() scheme (see XPointer xmlns() Scheme) which when prepended to a resolvable pointer allows for the definition of namespace prefixes to be used in XPaths in subsequent pointers. TEI Pointer schemes assume that un-prefixed element names in TEI Pointer XPaths are in the TEI namespace, `http://www.tei-c.org/ns/1.0`. The use of xmlns() is thus optional, provided no new prefixes need to be defined. If the schemes described here are used to address non-TEI elements, then any new prefixes to be used in pointer XPaths may be defined using the xmlns() scheme.

#### 17.2.4.2 xpath()

`Sequence xpath(XPATH)`

The xpath() scheme locates zero or more nodes within an XML Information Set. The single argument XPATH is an XPath selection pattern, as defined in XSLT 3.0, that returns a node or sequence of nodes. XPaths returning atomic values (e.g. substring()) are illegal in the xpath() scheme because they represent extracted values rather than locations in the source document. Because the schemes below involve starting at a node and navigating from there, and because attribute nodes have no intrinsic order, XPath expressions that address attribute nodes should be avoided in schemes other than xpath().

The example below, and all subsequent examples in this section refer to the following TEI fragment:

```
<div xml:lang="la" type="edition"
 xml:space="preserve"><ab>
<lb n="1" xml:id="line1"/><supplied reason="lost">si</supplied> non <choice><reg>habui</reg><or
<lb n="2"/>si<gap reason="illegible" quantity="3"
 unit="character"/>b<gap reason="illegible" quantity="3"
 unit="character"/>
 cohort<unclear>e</unclear> mi rescribas
<lb n="3"/><unclear>s</unclear>emp<unclear>er</unclear> in <choice><reg>mente</reg><orig>mentem</
 <choice><reg>habe</reg><orig>abe</orig></choice> supra res
<lb n="4"/>scriptas<gap reason="lost" extent="unknown"
 unit="character"/>
<lb n="5"/>auge et opto u<unclear>t</unclear> bene valeas</ab>
</div>
```

A TEI Pointer that referenced the <reg> element in the `choice` in line 1 of the example might look like: `#xpath(//lb[@n='1']/following-sibling::choice[1]/reg)`. Note that XPath values must be assumed to start from the document root. They cannot be relative to the element bearing the attribute that uses the pointer because TEI Pointers are URIs. Care should be taken to ensure that XPaths used in TEI Pointers match only a single node, unless multiple matches are desired. The examples that follow are relatively simple because the document they refer to is short and does not contain many elements.

When an XPath is interpreted by a TEI processor, the information set of the referenced document is interpreted without any additional information supplied by any schema processing that may or may not be present. In particular this means that no whitespace normalization is applied to a document before the XPath is interpreted.

This pointer scheme allows easy, direct use of the most widely-implemented XML query method. It is probably the most robust pointing mechanism for the common situation of selecting an XML element or its contents where an *xml:id* is not present. The ability to use element names and attribute names and values makes xpath() pointers more robust than the other mechanisms discussed in this section even if the designated document changes. For durability in the presence of editing, use of *xml:id* is always recommended when possible.

#### 17.2.4.3 left()

`Point left( IDREF | XPATH )`

The left() scheme locates the point immediately preceding the node addressed by its argument, which is either an XPATH as defined above or an IDREF, the value of an *xml:id* occurring in the document addressed by the base URI in effect for the pointer.

Example: the pointer `#left(//supplied[1])` indicates the point between the first `lb` and the first `supplied` in the example above.

Example: `#left(//gap[1])` indicates the point immediately before the first `gap` element in line two and the string `si`.

Example: `#left(line1)` indicates the point immediately before the <lb n=”1”> element.

#### 17.2.4.4 right()

`Point right( IDREF | XPATH )`

The right() scheme locates the point immediately following the node addressed by its argument.

Example: the pointer `#right(//lb[@n='3'])` indicates the point between the third `lb` and the `<unclear>s</unclear>` element in the example.

#### 17.2.4.5 string-index()

`Point string-index( IDREF | XPATH, OFFSET )`

The string-index() scheme locates a point based on character positions in a text stream relative to the node identified by the IDREF or XPATH parameter. The OFFSET parameter is a positive, negative, or zero integer which determines the position of the point. An offset of 0 represents the position immediately before the first character in either the first text node descendant of the node addressed in the first parameter or the first following text node, if the addressed element contains no text node descendants.

Example: `#string-index(//lb[@n='2'],1)` indicates the point between the 's' and the 'i' in the word 'si' in line 2.

**Note**: The OFFSET parameter (and similarly the LENGTH parameter found below in the string-range() scheme) are measured in characters. What is considered a single character will depend on the Normalization Form in use (see UNICODE NORMALIZATION FORMS). A letter followed by a combining diacritic counts as two characters, but the same diacritic precombined with a letter would count as a single character. Compare, for example, é (U+0060 followed by U+0301) and é (>U+00E9). These are equivalent, and a conversion between Normalization Forms C and D will transform one into the other. This specification does not mandate a particular Normalization Form (see *Precomposed and Combining Characters and Normalization*), but users and implementers should be aware that it affects the character count and therefore the result of evaluating pointers that rely on character counting.

#### 17.2.4.6 range()

`Sequence range( POINTER, POINTER[, POINTER, POINTER ...])`

The range() scheme takes as parameters one or more pairs of POINTERs, which are each members of the set IDREF, XPATH, left(), right(), or string-index(). A range() locates a (possibly non-contiguous) sequence beginning at the first POINTER parameter and ending at the last. If a POINTER locates a node (i.e. is an XPATH or IDREF), then that node is a member of the addressed sequence. If a sequence addressed by a range pointer overlaps, but does not wholly contain, an element (i.e. it contains only the start but not the end tag or vice-versa), then that element is not part of the sequence.

Range()s may address sequences of non-contiguous nodes. For example, a range() might select text beginning before an <app>, encompassing the content of a single <rdg> and continuing after the <app>.

Example: `#range(left(//lb[@n='3']),left(//lb[@n='4']))` indicates the whole of line 3 from the `<lb n="3"/>` to the point right before the following `<lb n="4"/>`.

Example: `#range(right(//lb[@n='3']),string-index(//lb[@n='3'],15))` indicates the sequence `<unclear>s</unclear>emp<unclear>er</unclear> in mente` (an element, a text node, another element, and a partial text node).

Example: `#range(string-index(//lb[@n='3'],7),string-index(//lb[@n='3'],10),string-index(` indicates the non-contiguous sequence 'in mentem'.

#### 17.2.4.7 string-range()

`Sequence string-range(IDREF | XPATH, OFFSET, LENGTH[, OFFSET, LENGTH ...])`

The string-range() scheme locates a sequence based on character positions in a text stream relative to the node identified by the first parameter. The location of the beginning of the addressed sequence is determined precisely as for string-index(). The OFFSET parameter is defined as above in string-index(). The LENGTH parameter is a positive integer that denotes the length of the text stream captured by the sequence. As with range(), the addressed sequence may contain text nodes and elements. The string-range() scheme can accept multiple OFFSET, LENGTH pairs to address a non-contiguous sequence in much the same way that range() can accept multiple pairs of pointers.

Because string-range() addresses points in the text stream, tags are invisible to it. For example, if an empty tag like <lb> is encountered while processing a string-range(), it will be included in the resulting sequence, but the LENGTH count will not increment when it is captured.

Example: `#string-range(//lb[@n='5'],0,27)` indicates the whole of line 5 from the text immediately following the `lb` to the point right before the closing `ab` tag.

Example: `#string-range(//lb[@n='3'],7,8)` indicates the sequence 'in mente'.

Example: `#string-range(//lb[@n='3'],7,3,15,6)` indicates the non-contiguous sequence 'in mentem'.

#### 17.2.4.8 match()

`Sequence match(IDREF | XPATH, 'REGEX' [, INDEX])`

The match scheme locates a sequence based on matching the REGEX parameter against a text stream relative to the reference node identified by the first parameter. REGEX is a regular expression as defined by XQuery 1.0 and XPath 2.0 Functions and Operators (Second Edition), with some modifications:

- Because the regular expression is delimited by apostrophe characters, any such characters (' or U+0027) occurring inside the expression must be escaped using the URI percent-encoding scheme `%27`.
- Regular expressions in `match()` are assumed to operate in single-line mode. The end of the string to be matched against is either the end of the text contained by the element in the first parameter or the end of the document, if that parameter indicates an empty element. The meta-character ^ therefore matches the beginning of the text stream inside or following the reference node, and the meta-character $ matches the end of that stream.

The optional INDEX parameter is an integer greater than 0 which specifies which match should be chosen when there is more than one possibility. If omitted, the first match in the text stream will be used to resolve the `match()`.

Like `string-range()`, `match()` may capture elements in the indicated sequence, even though they are ignored for purposes of evaluating the match.

Example: `#match(//lb[@n='5'],'opto.*valeas')` indicates the sequence `opto u<unclear>t</unclear> bene valeas` in line 5.

Example: `#match(//lb[@n='3'],'semper')` would indicate the word 'semper', but would not capture the `unclear` elements in `<unclear>s</unclear>emp<unclear>er</unclear>`, just their text children.

### 17.2.5 Canonical References

By 'canonical' reference we mean any means of pointing into documents, specific to a community or corpus. For example, biblical scholars might understand 'Matt 5:7' to mean 'the book called *Matthew*, chapter 5, verse 7.' They might then wish to translate the string 'Matt 5:7' into a pointer into a TEI-encoded document, selecting the element which corresponds to the seventh  element within the fifth  element within the  element with the *n* attribute valued 'Matt.'

Several elements in the TEI scheme (<gloss>, <ptr>, <ref>, and <term>) bear a special attribute, *cRef*, just for this purpose. Using the system described in this section, an encoder may specify references to canonical works in a discipline-familiar format, and expect software to derive a complete URI from it. The value of the *cRef* attribute is processed as described in this section, and the resulting URI reference is treated as if it were the value of the *target* attribute. The *cRef* and *target* attributes are mutually exclusive: only one or the other may be specified on any given occurrence of an element.

For the *cRef* attribute to function as required, a mechanism is needed to define the mapping between (for example) 'the book called *Matt*' and the part of the XML structure which corresponds with it. This is provided by the <refsDecl> element in the TEI header, which contains an algorithm for translating a canonical reference string (like Matt 5:7) into a URI such as `#xpath(//div[@n='Matt']/div[5]/div[7])`. The <refsDecl> element is described in section *2.3.6. The Reference System Declaration*; the following example is discussed in more detail below in section *17.2.5.1. Worked Example.* An alternative and less verbose method is described in section *17.2.5.4. Citation Structures.*

```
<refsDecl xml:id="biblical">
  <cRefPattern matchPattern="(.+) (.+):(.+)"
    replacementPattern="#xpath(//div[@n='$1']/div[@n='$2']/div[@n='$3]')">
    <p>This pointer pattern extracts and references the <q>book,</q>
      <q>chapter,</q> and <q>verse</q> parts of a biblical reference.</p>
  </cRefPattern>
  <cRefPattern matchPattern="(.+) (.+)"
    replacementPattern="#xpath(//div[@n='$1']/div[$2])">
    <p>This pointer pattern extracts and references the <q>book</q> and
    <q>chapter</q> parts of a biblical reference.</p>
  </cRefPattern>
  <cRefPattern matchPattern="(.+)"
    replacementPattern="#xpath(//div[@n='$1'])">
    <p>This pointer pattern extracts and references just the <q>book</q>
        part of a biblical reference.</p>
  </cRefPattern>
</refsDecl>
```

When an application encounters a canonical reference as the value of *cRef* attribute, it might follow this sequence of specific steps to transform it into a URI reference:

1. Ascertain the correct <refsDecl> following the rules summarized in section *16.3.3. Summary*.
2. For each <cRefPattern> element encountered in the appropriate <refsDecl>, in the order encountered:
   (a) match the value of the *cRef* attribute to the regular expression found as the value of the *matchPattern* attribute
   (b) if the value of the *cRef* attribute matches:
      i. take the value of the *replacementPattern* attribute and substitute the back references ($1, $2, etc.) with the corresponding matched substrings
      ii. the result is taken as if it were a relative or absolute URI reference specified on the *target* attribute; i.e., it should be used as is or combined with the current *xml:base* attribute value as usual
      iii. no further processing of this value of the *cRef* attribute against the <refsDecl> should take place
   (c) if, however, the value of the *cRef* attribute does not match the regular expression specified in the value of the *matchPattern* attribute, proceed to the next <cRefPattern>
3. If all the <cRefPattern> elements are examined in turn and none matches, the pointer fails.

The regular expression language used as the value of the *matchPattern* attribute is that used for the *pattern* facet of the World Wide Web Consortium's XML Schema Language in an Appendix to XML Schema Part 2.[99] The value of the *replacementPattern* attribute is simply a string, except that occurrences of '$1' through '$9' are replaced by the corresponding substring match. Note that since a maximum of nine substring matches are permitted, the string '$18' means 'the value of the first matched substring followed by the character '8'' as opposed to 'the eighteenth matched substring'. If there is a need for an actual string including a dollar sign followed by a digit that is not supposed to be replaced, the dollar sign should be written as `$$`. Implementations must convert `$$` to `$` during processing.

#### 17.2.5.1 Worked Example

Let us presume that with the example <refsDecl> above, an application comes across a *cRef* value of Matt 5:7. The application would first apply the regular expression `(.+) (.+):(.+)` to 'Matt 5:7'. This regular expression would successfully match. The first matched substring would be 'Matt', the second '5', and the third '7'. The application would then apply these substrings to the pattern `#xpath(//div[@n='$1']/div[$2]/div[$3])`, producing `#xpath(//div[@n='Matt']/div[5]/div[7])`.

If, however, the input string had been 'Matt 5', the first regular expression would not have matched. The application would have then tried the second, `(.+) (.+)`, producing a successful match, and the matched substrings 'Matt' and '5'. It would then have substituted those matched substrings into the pattern `#xpath(//div[@n='$1']/div[$2])` to produce a fragment identifier indicating the referenced element.

If the input string had been 'Matt', neither the first nor the second regular expressions would have successfully matched. The application would have then tried the third, `(.+)`, producing the matched substring 'Matt', and the URI Reference `#xpath(//div[@n='Matt'])`.

a <cRefPattern> should not reference more matched substrings. For example:

```
<cRefPattern  matchPattern="(.+) (.+):(.+)"
  replacementPattern="//div[@n='$1']/div[$2]/div[$3]/p[$4]"/>
```

is faulty, since only three matched substrings would have been produced, but a fourth (`$4`) was referenced.

#### 17.2.5.2 Complete and Partial URI Examples

In the above example, the value of *cRef* was used to generate a Fragment Identifier. An absolute URI could be generated directly, as in the following example.

[99] As always seems to be the case, no two regular expression languages are precisely the same. For those used to Perl regular expressions, be warned that while in Perl the pattern `tei` matches any string that contains *tei*, in the W3C language it only matches the string 'tei'.

```
<refsDecl  xml:id="USC">
  <cRefPattern  matchPattern="([0-9][0-9])\s*U\.?S\.?C\.?\s*[Cc](h(\.|ap(ter|\.)?)?)?\s*([1-9][0-9
    replacementPattern="http://uscode.house.gov/download/pls/$1C$5.txt">
    <p>Matches most standard references to particular
        chapters of the United States Code, e.g.
    <val>11USCC7</val>, <val>17 U.S.C. Chapter 3</val>, or
    <val>14 USC Ch. 5</val>. Note that a leading zero is
        required for the title (must be two digits), but is not
        permitted for the chapter number.</p>
  </cRefPattern>
  <cRefPattern  matchPattern="([0-9][0-9])\s*U\.?S\.?C\.?\s*[Pp](re(lim(inary)?)?)?\s*[Mm](at(eria
    replacementPattern="http://uscode.house.gov/download/pls/$1T.txt">
    <p>Matches references to the preliminary material for a
        given title, e.g. <val>11USCP</val>, <val>17 U.S.C.
           Prelim Mat</val>, or <val>14 USC pm</val>.</p>
  </cRefPattern>
  <cRefPattern  matchPattern="([0-9][0-9])\s*U\.?S\.?C\.?\s*[Aa](ppend(ix)?)?"
    replacementPattern="http://uscode.house.gov/download/pls/$1A.txt">
    <p>Matches references to the appendix of a given tile,
        e.g. <val>05USCA</val>, <val>11 U.S.C. Appendix</val>,
        or <val>18 USC Append</val>.</p>
  </cRefPattern>
</refsDecl>
<!-- ... -->
<p>The example in section 10 is taken
from <ref  cRef="17 USC Ch 1">Subject Matter and Scope of
    Copyright</ref>.</p>
```

See *17.2.3. Using Abbreviated Pointers* for another related use of the *matchPattern* and *replacementPattern* attributes.

#### 17.2.5.3 Miscellaneous Usages

Canonical reference pointers are intended for use by TEI encoders. However, this specification might be useful to the development of a process for recognizing canonical references in non-TEI documents (such as plain text documents), possibly as part of their conversion to TEI.

#### 17.2.5.4 Citation Structures

Citation structures provide a more thorough and concise mechanism for describing canonical references and the ways those references map on to parts of a TEI document. A <citeStructure> element describes a single step in a reference, such as 'Matt', and may nest to handle multi-part references.

The equivalent structure to the set of <cRefPattern>s in *17.2.5.1. Worked Example* would be:

```
<refsDecl  xml:id="biblical2">
  <citeStructure  unit="book"
    match="//body/div"  use="@n">
    <citeStructure  unit="chapter"  match="div"
      use="@n"  delim=" ">
      <citeStructure  unit="verse"  match="div"
        use="@n"  delim=":"/>
    </citeStructure>
  </citeStructure>
</refsDecl>
```

An application wishing to resolve a canonical reference such as Matt 5:7 might follow this procedure:

1. Ascertain the correct <refsDecl> following the rules summarized in section *16.3.3. Summary*.
2. Begin with the outer <citeStructure>. If it has a *delim* attribute and the reference begins with the value of *delim* then take the portion of the reference after the value of *delim* as input for the next child <citeStructure>.

3. For each nested <citeStructure>, if the input reference string contains the value of the *delim*, then split the string on the value of the *delim* attribute. If the input string does not contain *delim*, then stop. Take the portion of the input string after the value of *delim* and use it as the input string for the child <citeStructure>. After processing the outer <citeStructure>, the output will be `('Matt 5:7')`, after the second, `('Matt', '5:7')`. The end result will be a sequence like `('Matt','5','7')`.
4. For each item in the resulting sequence, resolve the matching node by evaluating the XPath in *match* with the predicate found in *use*, using the context of the previously matched node, if any. Start with the outer <citeStructure> and move to the next child <citeStructure> for each step in the sequence. For example, for the first <citeStructure>, we could construct an XPath `//div[@n='Matt']`. The full XPath after the reference Matt 5:7 is resolved will be `//div[@n='Matt']/div[@n='5']/div[@n='7']`.

One advantage <citeStructure> has is that it can be used to *generate* canonical references, using the declared citation structure to query the text structure. This means it is possible to automatically produce a list of resolvable citations for a TEI document. It also enables the automatic breaking of documents into smaller chunks for presentation and automated generation of tables of contents.

Citation structures may in addition specify how informational properties are to be extracted from the document sections they identify, using the <citeData> element. For example, if a TEI document is divided into chapters with a  per chapter and those chapters have titles, contained in <head> elements, then we might declare a citation structure for the document thus:

```
<citeStructure unit="chapter"
  match="/TEI/text/body/div" use="position()" delim="ch. ">
  <citeData property="http://purl.org/dc/terms/title"
    use="head"/>
</citeStructure>
```

This specifies that chapter references are given in the form 'ch. n', where 'n' is the position of the  in the <body> of the document, and that we may obtain the title of the chapter (identified by the Dublin Core property `title`) from the chapter heading. This would, for example, enable the automated generation of a chapter listing for the document.

## 17.3 Blocks, Segments, and Anchors

In this section, we discuss three general purposes elements which may be used to mark and categorize both a span of text and a point within one. These elements have several uses, most notably to provide elements which can be given identifiers for use when aligning or linking to parts of a document, as discussed elsewhere in this chapter. They also provide a convenient way of extending the semantics of the TEI markup scheme in a theory-neutral manner, by providing for two neutral or 'anonymous' elements to which the encoder can add any meaning not supplied by other TEI defined elements.

`<anchor>` (anchor point) attaches an identifier to a point within a text, whether or not it corresponds with a textual element.

`<ab>` (anonymous block) contains any component-level unit of text, acting as a container for phrase or inter level elements analogous to, but without the same constraints as, a paragraph.

`<seg>` (arbitrary segment) represents any segmentation of text below the 'chunk' level.

The elements <anchor>, <ab>, and <seg> are members of the class `att.typed`, from which they inherit the following attributes:

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.

`@subtype` (subtype) provides a sub-categorization of the element, if needed.

The elements <ab>, and <seg> are members of the class `att.fragmentable`, from which they inherit the following attribute:

`att.fragmentable` provides attributes for representing fragmentation of a structural element, typically as a consequence of some overlapping hierarchy.

@*part* specifies whether or not its parent element is fragmented in some way, typically by some other overlapping structure: for example a speech which is divided between two or more verse stanzas, a paragraph which is split across a page division, a verse line which is divided between two speakers.

The <seg> element is also a member of the class `att.segLike` from which it inherits the following attribute:

`att.segLike` provides attributes for elements used for arbitrary segmentation.

@*function* (function) characterizes the function of the segment.

The <anchor> element may be thought of as an empty <seg>, or as an artifice enabling an identifier to be attached to any position in a text. Like the <milestone> element discussed in section *3.11. Reference Systems*, it is useful where multiple views of a document are to be combined, for example, when a logical view based on paragraphs or verse lines is to be mapped on to a physical view based on manuscript lines. Like those elements, it is a member of the class `model.global` and can therefore appear anywhere within a document when the module defined by this chapter is included in a schema. Unlike the other elements in its class, the <anchor> element is primarily intended to mark an arbitrary point used for alignment, or as the target of a spanning element such as those discussed in section *12.3.1.4. Additions and Deletions*, rather than as a means of marking segment boundaries for some arbitrary segmentation of a text.

For example, suppose that we wish to mark the end of the fifth word following each occurrence of some term in a particular text, perhaps to assist with some collocational analysis. This can most easily be done with the help of the <anchor> element, as follows:

```
English language. Except for not very<anchor xml:id="eng1"/>
English at all at the time<anchor xml:id="eng2"/>
English was still full of flaws<anchor xml:id="eng3"/>
English. This was revised by young
<anchor xml:id="eng4"/>
```

In section *17.5.1. Correspondence* we discuss ways in which these <anchor> points might be used to represent an alignment such as one might get in a keyword-in-context concordance.

The <seg> element may be used at the encoder's discretion to mark almost any segment of the text of interest for processing. One use of the element is to mark text features for which no appropriate markup is otherwise defined, i.e. as a simple extension mechanism. Another use is to provide an identifier for some segment which is to be pointed at by some other element, i.e. to provide a target, or a part of a target, for a <ptr> or other similar element.

Several examples of uses for the <seg> element are provided elsewhere in these Guidelines. For example:

- as a means of marking segments significant in a metrical or rhyming analysis (see section *6.4. Rhyme and Metrical Analysis*)
- as a means of marking typographic lines in drama (see section *7.2. The Body of a Performance Text*) or title pages (see section *4.6. Title Pages*)
- as a means of marking prosody- or pause-defined units in transcribed speech (see section *8.4.1. Segmentation*)
- as a means of marking linguistic or other analyses in a theory-neutral manner (see chapter *18. Simple Analytic Mechanisms* passim)

In the following simple example, the <seg> element simply delimits the extent of a stutter, a textual feature for which no element is provided in these Guidelines.

```
<q>Don't say <q>
   <seg type="stutter">I-I-I</seg>'m afraid,</q> Melvin, just say <q>I'm
    afraid.</q>
</q>
```

The <seg> element is particularly useful for the markup of linguistically significant constituents such as the phrases that may be the output of an automatic parsing system. This example also demonstrates the use of the *xml:id* attribute to carry an identifier which other parts of a document may use to point to, or align with:

```
<seg xml:id="bl0034" type="sentence">
  <seg xml:id="bl0034.1" type="phrase">Literate and illiterate speech</seg>
  <seg xml:id="bl0034.2" type="phrase">in a language like English</seg>
  <seg xml:id="bl0034.3" type="phrase">are plainly different.</seg>
</seg>
```

As the above example shows, <seg> elements may be nested directly within one another, to any degree of analysis considered appropriate. This is taken a little further in the following example, where the *type* and *subtype* attributes have been used to further categorize each word of the sentence (the *xml:id* attributes have been removed to reduce the complexity of the example):

```
<seg type="sentence" subtype="declarative">
  <seg type="phrase" subtype="noun">
    <seg type="word" subtype="adjective">Literate</seg>
    <seg type="word" subtype="conjunction">and</seg>
    <seg type="word" subtype="adjective">illiterate</seg>
    <seg type="word" subtype="noun">speech</seg>
  </seg>
  <seg type="phrase" subtype="preposition">
    <seg type="word" subtype="preposition">in</seg>
    <seg type="word" subtype="article">a</seg>
    <seg type="word" subtype="noun">language</seg>
    <seg type="word" subtype="preposition">like</seg>
    <seg type="word" subtype="noun">English</seg>
  </seg>
  <seg type="phrase" subtype="verb">
    <seg type="word" subtype="verb">are</seg>
    <seg type="word" subtype="adverb">plainly</seg>
    <seg type="word" subtype="adjective">different</seg>
  </seg>
  <seg type="punct">.</seg>
</seg>
```

(The example values shown are chosen for simplicity of comprehension, rather than verisimilitude). It should also be noted that specialized segment elements are defined in section *18.1. Linguistic Segment Categories* to facilitate this particular kind of analysis. These allow for the explicit markup of units called *s-units*, *clauses*, *phrases*, *words*, *morphemes*, and *characters*, which may be felt preferable to the more generic approach typified by use of the <seg> element. Using these, the first phrase above might be encoded simply as

```
<phr type="noun">
  <w type="adjective">Literate</w>
  <w type="conjunction">and</w>
  <w type="adjective">illiterate</w>
  <w type="noun">speech</w>
</phr>
```

Note the way in which the *type* attribute of these specialized elements now carries the value carried by the *subtype* attribute of the more general <seg> element. For an analysis not using these traditional linguistic categories however, the <seg> element provides a simple but powerful mechanism.

In language corpora and similar material, the <seg> element may be used to provide an end-to-end segmentation as an alternative to the more specific <s> element proposed in chapter *18.1. Linguistic Segment Categories* for the markup of orthographic sentences, or *s-units*. However, it may be more useful to use the <s> element for this purpose, since

this means that the <seg> element can then be used to mark both features within s-units and segments composed of s-units, as in the following example:[100]

```
<seg  xml:id="s1s3"  type="narrative_unit">
  <s  xml:id="s1">Sigmund, the <seg  type="patronymic">son of Volsung</seg>,
     was a king in Frankish country.</s>
  <s  xml:id="s2">Sinfiotli was the eldest of his sons.</s>
  <s  xml:id="s3"> ... </s>
</seg>
```

Like other elements, the <seg> tag must be properly enclosed within other elements. Thus, a single <seg> element can be used to group together words in different sentences only if the sentences are not themselves tagged. The first of the following two encodings is legal, but the second is not.

```
Give me <seg  type="phrase">a dozen. Or two or three.</seg>
```

```
  <!-- Illegal! -->
<s>Give me <seg type="phrase">a dozen.</s>
<s>Or two or three.</s></seg>
```

The *part* attribute may be used as one simple method of overcoming this restriction:

```
<s>Give me <seg  type="phrase"  part="I">a dozen.</seg>
</s>
<s>
  <seg  part="F">Or two or three.</seg>
</s>
```

Another solution is to use the <join> element discussed in section *17.7. Aggregation*; this requires that each of the <seg> elements be given an identifier. For further discussion of this generic encoding problem, see also chapter *21. Non-hierarchical Structures*.

The <seg> element has the same content as a paragraph in prose: it can therefore be used to group together consecutive sequences of `model.inter` class elements, such as lists, quotations, notes, stage directions, etc. as well as to contain sequences of phrase-level elements. It cannot however be used to group together sequences of paragraphs or similar text units such as verse lines; for this purpose, the encoder should use intermediate pointers, as described in section *17.1.4. Intermediate Pointers* or the methods described in section *17.7. Aggregation*. It is particularly important that the encoder provide a clear description of the principles by which a text has been segmented, and the way in which that segmentation is represented. This should include a description of the method used and the significance of any categorization codes. The description should be provided as a series of paragraphs within the <segmentation> element of the encoding description in the TEI header, as described in section *2.3.3. The Editorial Practices Declaration*.

The <seg> element may also be used to encode simultaneous or mutually exclusive variants of a text when the more special purpose elements for simple editorial changes, abbreviation and expansion, addition and deletion, or for a critical apparatus are not appropriate. In these circumstances, one <seg> is encoded for each possible variant, and the set of them is enclosed in a <choice> element.

For example, if one were writing dual-platform instructions for installation of software, it might be useful to use <seg> to record platform-specific pieces of mutually exclusive text.

[100] See section *18.3. Spans and Interpretations*, where the text from which this fragment is taken is analyzed.

```
…pressing <choice>
  <seg  type="platform"  subtype="Mac">option</seg>
  <seg  type="platform"  subtype="PC">alt</seg>
</choice>-f will …
```

Elsewhere in this chapter we provide a number of examples where the <seg> element is used simply to provide an element to which an identifier may be attached, for example so that another segment may be linked or related to it in some way.

The <ab> (anonymous block) element functions like the <seg> element, but is used for portions of the text which occur at the *chunk* level (See *1.3.2.1. Informal Element Classifications*). It is therefore a member of the `model.pLike` class. Like <seg>, but unlike <p>, <ab> may nest within other <ab> elements. So it may be used to represent texts with irregular *chunk*-level structures.

Probably the most common use for <ab> is as a container for text that does not naturally belong in a paragraph. The text of an epitaph, for example, may not be paragraph-like, but still requires a container:

```
<ab  style="text-align: center;">
  <lb/>SACRED
<lb/>to the memory of
<lb/>
  <abbr>M<hi  rend="superscript">rs</hi>
    <am>.</am>
  </abbr> ALICE HUTTON
<lb/>born at Che<g  ref="#long-s">s</g>terfield in
<lb/>Great Britain,
<lb/>
  <abbr>Jan<hi  rend="superscript">y</hi>
    <am>.</am>
  </abbr>
  <abbr>25<hi  rend="superscript">th</hi>
    <am>.</am>
  </abbr> 1722:
<lb/>and died in Carolina
<lb/>
  <abbr>Sept<hi  rend="superscript">r</hi>
    <am>.</am>
  </abbr>
  <abbr>26<hi  rend="superscript">th</hi>
    <am>.</am>
  </abbr> 1777.
<lb/>This <g  ref="#long-s">s</g>tone
<lb/>is placed by her only Child
<lb/>ANN BOOTH POLLOK
</ab>
```

Encoders may also prefer to treat other types of short documents, such as receipts, in this way.

The <ab> element may also be used, for example, to tag the canonical verse divisions of Biblical texts:

```
<div1  n="Gen"  type="book">
  <head>The First Book of Moses, Called</head>
  <head  type="main">Genesis</head>
  <div2  n="1"  type="chapter">
    <ab  n="1">In the beginning God created the heaven and the
        earth.</ab>
    <ab  n="2">And the earth was without form, and void; and darkness
```

```
    <hi>was</hi> upon the face of the deep. And the Spirit of God
        moved upon the face of the waters.</ab>
    <ab  n="3">And God said, Let there be light: and there was
        light.</ab>
  </div2>
</div1>
```

In other cases, where the text clearly indicates paragraph divisions containing one or more verses, the <p> element may be used to tag the paragraphs, and the <seg> element used to subdivide them. The <ab> element is provided as an alternative to the <p> element; it may *not* be used within paragraphs. The <seg> element, by contrast, may appear only within and not between paragraphs (or anonymous block elements).

```
<div1  n="Gen"  type="book">
  <head>Das Erste Buch Mose.</head>
  <div2  n="1"  type="chapter">
    <p>
      <seg  n="1">Am Anfang schuff Gott Himel vnd Erden.</seg>
      <seg  n="2">Vnd die Erde war wüst vnd leer / vnd es war
            finster auff der Tieffe / Vnd der Geist Gottes schwebet auff
            dem Wasser.</seg>
    </p>
    <p>
      <seg  n="3">Vnd Gott sprach / Es werde Liecht / Vnd es ward
            Liecht.</seg>
    </p>
  </div2>
</div1>
```

The <ab> element is also useful for marking dramatic speeches when it is not clear whether the speech is to be regarded as prose or verse. If, for example, an encoder does not wish to express an opinion as to whether the opening lines of Shakespeare's *The Tempest* are to be regarded as prose or as verse, they might be tagged as follows:

```
<div1  n="I"  type="act">
  <div2  n="1"  type="scene">
    <head  rend="italic">Actus primus, Scena prima.</head>
    <stage  rend="italic"  type="setting"> A tempestuous noise of
        Thunder and Lightning heard:
        Enter a Ship-master, and a Boteswaine.</stage>
    <sp>
      <speaker>Master.</speaker>
      <ab>Bote-swaine.</ab>
    </sp>
    <sp>
      <speaker>Botes.</speaker>
      <ab>Heere Master: What cheere?</ab>
    </sp>
    <sp>
      <speaker>Mast.</speaker>
      <ab>Good: Speake to th' Mariners: fall too't, yarely,
            or we run our selues a ground, bestirre, bestirre.
      <stage  type="move">Exit.</stage>
      </ab>
    </sp>
    <stage  type="move">Enter Mariners.</stage>
```

```
    <sp>
      <speaker>Botes.</speaker>
      <ab>Heigh my hearts, cheerely, cheerely my harts: yare, yare:
            Take in the toppe-sale: Tend to th' Masters whistle: Blow
            till thou burst thy winde, if roome e-nough.</ab>
    </sp>
  </div2>
</div1>
```

See further *3.13.2. Core Tags for Drama* and *7.2.5. Speech Contents*.

## 17.4 Synchronization

In the previous section we discussed two particular kinds of alignment: alignment of parallel texts in different languages; and alignment of texts and portions of an image. In this section we address another specialized form of alignment: synchronization. The need to mark the relative positions of text components with respect to time arises most naturally and frequently in transcribed spoken texts, but it may arise in any text in which quoted speech occurs, or events are described within a time frame. The methods described here are also generalizable for other kinds of alignment (for example, alignment of text elements with respect to space).

### 17.4.1 Aligning Synchronous Events

Provided that explicit elements are available to represent the parts or places to be synchronized, then the global linking attribute *synch* may be used to encode such synchronization, once it has been identified.

`att.global.linking` provides a set of attributes for hypertextual linking.

`@synch` (synchronous) points to elements that are synchronous with the current element.

This is another of the attributes made globally available by the mechanism described in the introduction to this chapter. Alternatively, the <link> and <linkGrp> elements may be used to make explicit the fact that the synchronous elements are aligned.

To illustrate the use of these mechanisms for marking synchrony, consider the following representation of a spoken text:

```
  B: The first time in twenty five years, we've cooked Christmas
(unclear) for a blooming great load of people.
A: So you're [1] (unclear) [2]
B: [1] It will be [2] nice in a way, but, [3] be strange. [4]
A: [3] Yeah [4], yeah, cos it, it's [5] the [6]
B: [5] not [6]
```

This representation uses numbers in brackets to mark the points at which speakers overlap each other. For example, the *[1]* in A's first speech is to be understood as coinciding with the *[1]* in B's second speech.[101]

To encode this we use the spoken texts module, described in chapter *8. Transcriptions of Speech*, together with the module described in the present chapter. First, we transcribe this text, marking the synchronous points with <anchor> elements, and providing a *synch* attribute on one of each of the pairs of synchronous anchors. As noted in the example given above (section *17.5.2. Alignment of Parallel Texts*), correspondence, and hence synchrony, is a symmetric relation; therefore the attribute need only be specified on one of the pairs of synchronous anchors.

```
<div  xml:id="BNC-d1"  type="convers">
  <u  xml:id="u2b"  who="#b"> The first time in twenty five years,
     we've cooked Christmas <unclear> for a blooming great
```

[101] This sample is taken from a conversation collected and transcribed for the British National Corpus.

```
        load of people.</unclear>
  </u>
  <u  xml:id="u3a"  who="#a">So you're
  <anchor  synch="#t1b"  xml:id="t1a"/>
    <unclear>
      <anchor  synch="#t2b"  xml:id="t2a"/>
    </unclear>
  </u>
  <u  xml:id="u3b"  who="#b">
    <anchor  xml:id="t1b"/>It will be <anchor  xml:id="t2b"/>
     nice in a way, but, <anchor  xml:id="t3b"/>
     be strange.<anchor  xml:id="t4b"/>
  </u>
  <u  xml:id="u4a"  who="#a">
    <anchor  synch="#t3b"  xml:id="t3a"/>Yeah
  <anchor  synch="#t4b"  xml:id="t4a"/>, yeah, cos it, its
  <anchor  synch="#t5b"  xml:id="t5a"/>the
  <anchor  synch="#t6b"  xml:id="t6a"/>
  </u>
  <u  xml:id="u4b"  who="#b">
    <anchor  xml:id="t5b"/>not<anchor  xml:id="t6b"/>
  </u>
<!-- ... -->
</div>
```

We can encode this same example using <link> and <linkGrp> elements to make the temporal alignment explicit. A <back> element has been used to enclose the <linkGrp> element, but the links may be located anywhere the encoder finds convenient:

```
<back>
  <linkGrp  xml:id="lg1"
    domains="#BNC-d1 #BNC-d1"  targFunc="speaker.a speaker.b"
    type="synchronous_alignment">
    <link  xml:id="L1"  target="#t1a #t1b"/>
    <link  xml:id="L2"  target="#t2a #t2b"/>
    <link  xml:id="L3"  target="#t3a #t3b"/>
    <link  xml:id="l4"  target="#t4a #t4b"/>
    <link  xml:id="l5"  target="#t5a #t5b"/>
    <link  xml:id="l6"  target="#t6a #t6b"/>
  </linkGrp>
</back>
```

The *xml:id* attributes are provided for the <link> and <linkGrp> elements here for reasons discussed in the next section, *17.4.2. Placing Synchronous Events in Time*.

As with other forms of alignment, synchronization may be expressed between stretches of speech as well as between points. When complete utterances are synchronous, for example, if one person says *What?* and another *No!* at the same time, that can be represented without <anchor> elements as follows.

```
<u  synch="#u02"  xml:id="u01"  who="#a">What?</u>
<u  xml:id="u02"  who="#b">No!</u>
```

A simple way of expressing *overlap* (where one speaker starts speaking before another has finished) is thus to use the <seg> element to encode the overlapping portions of speech. For example,

```
<u  who="#a"> So you're <unclear  synch="#u-b1"/>
</u>
<u  who="#b">
  <seg  xml:id="u-b1"> It will be </seg> nice in a way, but,
<seg  synch="#u-a3"> be strange. </seg>
</u>
<u  who="#a">
  <seg  xml:id="u-a3"> Yeah </seg>, yeah, cos it,
its <seg  synch="#u-b2"> the </seg>
</u>
<u  xml:id="u-b2"  who="#b"> not </u>
```

Note in this encoding how synchronization has been effected between an empty <unclear> element and the content of a <seg> element, and between the content of a <u> element and that of another <seg>, using the *synch* attribute. Alternatively, a <linkGrp> could be used in the same way as above.

### 17.4.2 Placing Synchronous Events in Time

A synchronous alignment specifies which points in a spoken text occur at the same time, and the order in which they occur, but does not say at what time those points actually occur. If that information is available to the encoder it can be represented by means of the <when> and <timeline> elements, whose description and attributes are the following:

`<when>` indicates a point in time either relative to other elements in the same timeline tag, or absolutely.

- `@absolute` supplies an absolute value for the time.
- `@interval` specifies a time interval either as a number or as one of the keywords defined by the datatype `teidata.interval`.
- `@unit` specifies the unit of time in which the *interval* value is expressed, if this is not inherited from the parent <timeline>.
- `@since` identifies the reference point for determining the time of the current <when> element, which is obtained by adding the interval to the time of the reference point.

`<timeline>` (timeline) provides a set of ordered points in time which can be linked to elements of a spoken text to create a temporal alignment of that text.

- `@origin` designates the origin of the timeline, i.e. the time at which it begins.
- `@interval` specifies a time interval either as a positive integral value or using one of a set of predefined codes.
- `@unit` specifies the unit of time corresponding to the *interval* value of the timeline or of its constituent points in time.

Each <when> element indicates a point in time, either directly by means of the *absolute* attribute, whose value is a string which specifies a particular time, or indirectly by means of the *since* attribute, which points to another <when>. If the *since* is used, then the *interval* and *unit* attributes should also be used to indicate the amount of time that has elapsed since the time specified by the element pointed to by the *since* attribute; the value -1 can be given to indicate that the interval is unknown.

If the <when> elements are uniformly spaced in time, then the *interval* and *unit* values need be given once in the <timeline>, and not repeated in any of the <when> elements. If the intervals vary, but the units are all the same, then the *unit* attribute alone can be given in the <timeline> element, and the *interval* attribute given in the <when> element.

The *origin* attribute in the <timeline> element points to a <when> element which specifies the reference or origin for the timings within the <timeline>; this must, of course, specify its position in time absolutely. If the origin of a timeline is unknown, then this attribute may be omitted.

The following <timeline> might be used to accompany the marked up conversation shown in the preceding section:

```
<timeline xml:id="tL1" origin="#w0"
  unit="ms">
  <when xml:id="w0" absolute="11:30:00"/>
  <when xml:id="w1" interval="unknown"
    since="#w0"/>
  <when xml:id="w2" interval="100"
    since="#w1"/>
  <when xml:id="w3" interval="200"
    since="#w2"/>
  <when xml:id="w4" interval="150"
    since="#w3"/>
  <when xml:id="w5" interval="250"
    since="#w4"/>
  <when xml:id="w6" interval="100"
    since="#w5"/>
</timeline>
```

The information in this <timeline> could now be linked to the information in the <linkGrp> which provides the temporal alignment (synchronization) for the text, as follows:

```
<linkGrp type="temporal_specification"
  domains="#lg1 #tL1" targFunc="synch.points when">
  <link target="#L1 #w1"/>
  <link target="#L2 #w2"/>
  <link target="#L3 #w3"/>
  <link target="#l4 #w4"/>
  <link target="#l5 #w5"/>
  <link target="#l6 #w6"/>
</linkGrp>
```

To avoid the need for two distinct link groups (one marking the synchronization of anchors with each other, and the other marking their alignment with points on the time line) it would be better to link the <when> elements with the synchronous points directly:

```
<linkGrp type="temporal_specification"
  domains="#BNC-d1 #BNC-d1 #tL1" targFunc="speaker.a speaker.b when">
  <link target="#t1a #t1b #w1"/>
  <link target="#t2a #t2b #w2"/>
  <link target="#t3a #t3b #w3"/>
  <link target="#t4a #t4b #w4"/>
  <link target="#t5a #t5b #w5"/>
  <link target="#t6a #t6b #w6"/>
</linkGrp>
```

Finally, suppose that a digitized audio recording is also available, and an XML file that assigns identifiers to the various temporal spans of sound is available. For example, the following Synchronized Multimedia Integration Language (SMIL, pronounced ”smile”) fragment:

```
<audio xmlns="http://www.w3.org/2001/SMIL20/Language" src="rtsp://soundstage.pi.cnr.it:554/home/a
  xml:id="au1" begin="05.2s"/>
<audio xmlns="http://www.w3.org/2001/SMIL20/Language" src="rtsp://soundstage.pi.cnr.it:554/home/a
  xml:id="au2" begin="05.7s"/>
```

```
<audio xmlns="http://www.w3.org/2001/SMIL20/Language"  src="rtsp://soundstage.pi.cnr.it:554/home/a
  xml:id="au3"  begin="05.9s"/>
<audio xmlns="http://www.w3.org/2001/SMIL20/Language"  src="rtsp://soundstage.pi.cnr.it:554/home/a
  xml:id="au4"  begin="06.3s"/>
<audio xmlns="http://www.w3.org/2001/SMIL20/Language"  src="rtsp://soundstage.pi.cnr.it:554/home/a
  xml:id="au5"  begin="06.9s"/>
<audio xmlns="http://www.w3.org/2001/SMIL20/Language"  src="rtsp://soundstage.pi.cnr.it:554/home/a
  xml:id="au6"  begin="07.4s"/>
```

URIs pointing to the <audio> elements could also be included as a fourth component in each of the above <link> elements, thus providing a synchronized audio track to complement the transcribed text.

For further discussion of this and related aspects of encoding transcribed speech, refer to chapter *8. Transcriptions of Speech*.

## 17.5 Correspondence and Alignment

In this section we introduce the notions of *correspondence*, expressed by the *corresp* attribute, and of *alignment*, which is a special kind of correspondence involving an ordered set of correspondences. Both cases may be represented using the <link> and <linkGrp> elements introduced in section *17.1. Links*. We also discuss the special case of alignment in time or *synchronization*, for which special purpose elements are proposed in section *17.4. Synchronization*.

### 17.5.1 Correspondence

A common requirement in text analysis is to represent correspondences between two or more parts of a single document, or between places in different documents. Provided that explicit elements are available to represent the parts or places to be linked, then the global linking attribute *corresp* may be used to encode such correspondence, once it has been identified.

att.global.linking provides a set of attributes for hypertextual linking.

@*corresp* (corresponds) points to elements that correspond to the current element in some way.

This is one of the attributes made available by the mechanism described in the introduction to this chapter (*17. Linking, Segmentation, and Alignment*). Correspondence can also be expressed by means of the <link> element introduced in section *17.1. Links*.

Where the correspondence is between *spans*, the <seg> element should be used, if no other element is available. Where the correspondence is between *points*, the <anchor> element should be used, if no other element is available.

The use of the *corresp* attribute with spans of content is illustrated by the following example:

```
<title  xml:id="SHIRLEY">Shirley</title>, which made
its Friday night debut only a month ago, was
not listed on <name  xml:id="NBC">NBC</name>'s new schedule,
although <seg  corresp="#NBC"  xml:id="NETWORK">the network</seg>
says <seg  corresp="#SHIRLEY"  xml:id="SHOW">the show</seg>
still is being considered.
```

Here the anaphoric phrases *the network* and *the show* have been associated directly with the elements to which they refer by means of *corresp* attributes. This mechanism is simple to apply, but has the drawback that it is not possible to specify more exactly what kind of correspondence is intended. Where this attribute is used, therefore, encoders are encouraged to specify their intent in the associated encoding description in the TEI header.

Essentially, what the *corresp* attribute does is to specify that elements bearing this attribute and those to which the attribute points are doubly linked. In the example above, the use of the *corresp* attribute indicates that the <seg> element containing 'the show' and the <title> element containing 'Shirley' correspond to each other: the correspondence relationship is not 'from' one to the other, but 'between' the two objects. It is thus different from the *target* attribute, and provides functionality more similar to that of the <link> and <linkGrp> elements defined in section *17.1. Links*, although it lacks the ability to indicate more precisely what kind of correspondence is intended as in the following retagging of the preceding example.

```
<title  xml:id="shirley">Shirley</title>, which made
its Friday night debut only a month ago, was not
listed on <name  xml:id="nbc">NBC</name>'s new schedule,
although <seg  xml:id="network">the network</seg> says
<seg  xml:id="show">the show</seg> still is being considered.

<linkGrp  type="anaphoric_link"
  targFunc="antecedent anaphor">
  <link  target="#shirley #show"/>
  <link  target="#nbc #network"/>
</linkGrp>
```

In the following example, we use the same mechanism to express a correspondence amongst the anchors introduced following the fifth word after *English* in a text:

```
English language. Except for not very<anchor  xml:id="en1"/>
<!-- ... -->
English at all at the time
<anchor  xml:id="en2"/>
<!-- ... -->
English was still full of flaws
<anchor  xml:id="en3"/>
<!-- ... -->
English. This was revised by young
<anchor  xml:id="en4"/>
<!-- ... -->
<linkGrp  type="five-word_collocates">
  <link  type="collocates_of_ENGLISH"
    target="#en1 #en2 #en3 #en4"/>
<!-- ... -->
</linkGrp>
```

### 17.5.2 Alignment of Parallel Texts

One very important application area for the alignment of parallel texts is multilingual corpora. Consider, for example, the need to align 'translation pairs' of sentences drawn from a corpus such as the Canadian Hansard, in which each sentence is given in both English and French. Concerning this problem, Gale and Church write:

> Most English sentences match exactly one French sentence, but it is possible for an English sentence to match two or more French sentences. The first two English sentences [in the example below] illustrate a particularly hard case where two English sentences align to two French sentences. No smaller alignments are possible because the clause '…sales…were higher…' in the first English sentence corresponds to (part of) the second French sentence. The next two alignments … illustrate the more typical case where one English sentence aligns with exactly one French sentence. The final alignment matches two English sentences to a single French sentence. These alignments [which were produced by a computer program] agreed with the results produced by a human judge.[102]

The alignment produced by Gale and Church's program can be expressed in four different ways. The encoder must first decide whether to represent the alignment in terms of points within each text (using the <anchor> element) or in terms of whole stretches of text, using the <seg> element. To some extent the choice will depend on the process by which the software works out where alignment occurs, and the intention of the encoder. Secondly, the encoder may elect to represent the actual encoding using either *corresp* attributes attached to the individual <anchor> or <seg> elements, or using a free-standing <linkGrp> element.

We present first a solution using <anchor> elements bearing only *corresp* attributes:

[102] See ?, from which the example in the text is taken.

```
<div xml:lang="en" type="subsection">
  <p>
    <anchor corresp="#fa1" xml:id="ea1"/>According to our survey, 1988
     sales of mineral water and soft drinks were much higher than in 1987,
     reflecting the growing popularity of these products. Cola drink
     manufacturers in particular achieved above-average growth rates.
  <anchor corresp="#fa2" xml:id="ea2"/>The higher turnover was largely
     due to an increase in the sales volume.
  <anchor corresp="#fa3" xml:id="ea3"/>Employment and investment levels also climbed.
  <anchor corresp="#fa4" xml:id="ea4"/>Following a two-year transitional period,
     the new Foodstuffs Ordinance for Mineral Water came into effect on
     April 1, 1988. Specifically, it contains more stringent requirements
     regarding quality consistency and purity guarantees.</p>
</div>
<div xml:lang="fr" type="subsection">
  <p>
    <anchor corresp="#ea1" xml:id="fa1"/>Quant aux eaux minérales
     et aux limonades, elles rencontrent toujours plus d'adeptes. En effet,
     notre sondage fait ressortir des ventes nettement supérieures
     à celles de 1987, pour les boissons à base de cola
     notamment. <anchor corresp="#ea2" xml:id="fa2"/>La progression des
     chiffres d'affaires résulte en grande partie de l'accroissement
     du volume des ventes. <anchor corresp="#ea3" xml:id="fa3"/>L'emploi et
     les investissements ont également augmenté.
  <anchor corresp="#ea4" xml:id="fa4"/>La nouvelle ordonnance fédérale
     sur les denrées alimentaires concernant entre autres les eaux
     minérales, entrée en vigueur le 1er avril 1988 après
     une période transitoire de deux ans, exige surtout une plus
     grande constance dans la qualité et une garantie de la
     pureté.</p>
</div>
```

There is no requirement that the *corresp* attribute be specified in both English and French texts, since (as noted above) this attribute is defined as representing a mutual association. However, it may simplify processing to do so, and also avoids giving the impression that the English is translating the French, or vice versa. More seriously, this encoding does not make explicit that it is in fact the entire stretch of text between the anchors which is being aligned, not simply the points themselves. If for example one text contained material omitted from the other, this approach would not be appropriate.

We now present the same passage using the alternative <linkGrp> mechanism and marking explicitly the segments which have been aligned:

```
<div xml:id="div-e" xml:lang="en"
  type="subsection">
  <p>
    <seg xml:id="e_1">According to our survey, 1988 sales of mineral
        water and soft drinks were much higher than in 1987,
        reflecting the growing popularity of these products. Cola
        drink manufacturers in particular achieved above-average
        growth rates.</seg>
    <seg xml:id="e_2">The higher turnover was largely due to an
        increase in the sales volume.</seg>
    <seg xml:id="e_3">Employment and investment levels also climbed.</seg>
    <seg xml:id="e_4">Following a two-year transitional period, the new
        Foodstuffs Ordinance for Mineral Water came into effect on
        April 1, 1988. Specifically, it contains more stringent
```

```
        requirements regarding quality consistency and purity
        guarantees.</seg>
  </p>
</div>
<div  xml:id="div-f"  xml:lang="fr"
  type="subsection">
  <p>
    <seg  xml:id="f_1">Quant aux eaux minérales et aux limonades,
        elles rencontrent toujours plus d'adeptes. En effet, notre
        sondage fait ressortir des ventes nettement
        supérieures à celles de 1987, pour les
        boissons à base de cola notamment.</seg>
    <seg  xml:id="f_2">La progression des chiffres d'affaires
        résulte en grande partie de l'accroissement du volume
        des ventes.</seg>
    <seg  xml:id="f_3">L'emploi et les investissements ont
        également augmenté.</seg>
    <seg  xml:id="f_4">La nouvelle ordonnance fédérale sur
        les denrées alimentaires concernant entre autres les
        eaux minérales, entrée en vigueur le 1er avril
        1988 après une période transitoire de deux
        ans, exige surtout une plus grande constance dans la
        qualité et une garantie de la pureté.</seg>
  </p>
</div>
<linkGrp  type="alignment"
  domains="#div-e #div-f">
  <link  target="#e_1 #f_1"/>
  <link  target="#e_2 #f_2"/>
  <link  target="#e_3 #f_3"/>
  <link  target="#e_4 #f_4"/>
</linkGrp>
```

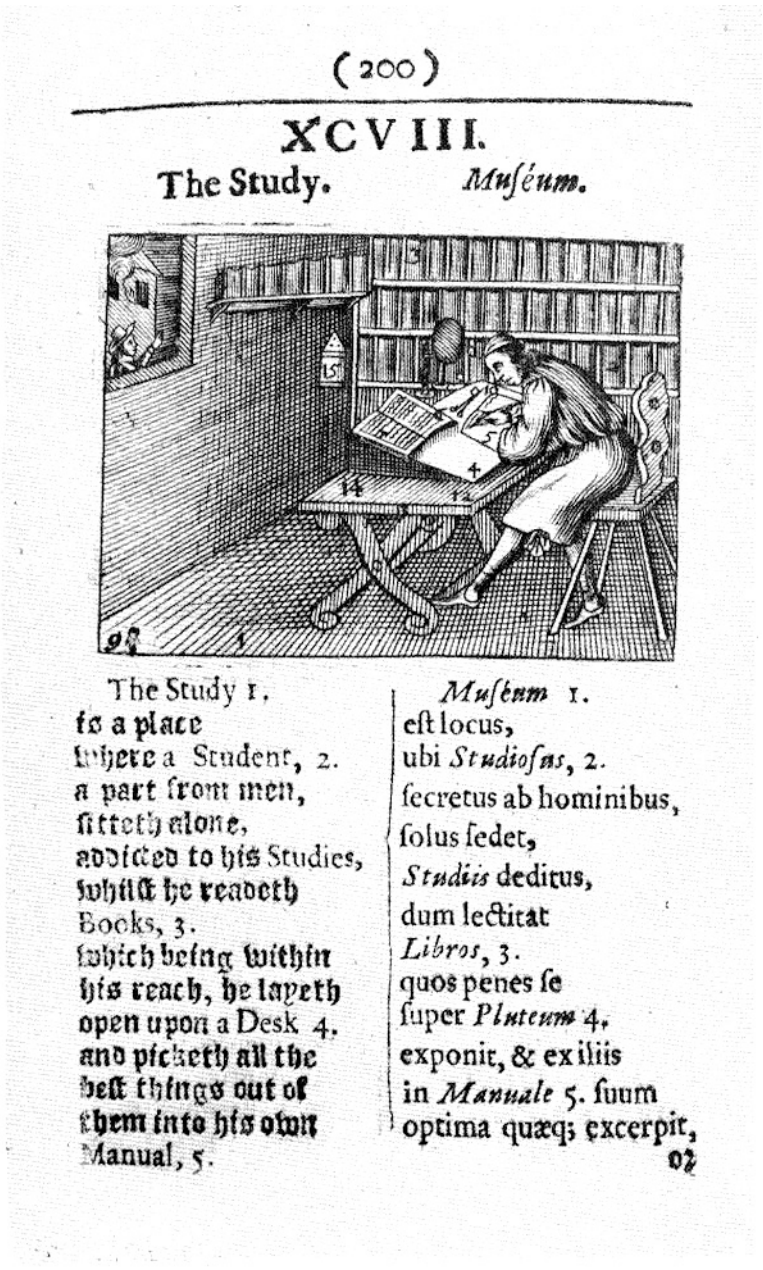
Note that use of the <ab> element allows us to mark up the orthographic sentences in both languages independently of the alignment: the first translation pair in this example might be marked up as follows:

```
<div  xml:id="english"  xml:lang="en"
  type="subsection">
  <ab  xml:id="english1">
    <s>According to our survey, 1988 sales of mineral water and soft
        drinks were much higher than in 1987, reflecting the growing popularity
        of these products.</s>
    <s>Cola drink manufacturers in particular achieved above-average
        growth rates.</s>
  </ab>
</div>
<div  xml:id="french"  xml:lang="fr"
  type="subsection">
  <ab  xml:id="french1">
    <s  xml:id="fs1">Quant aux eaux minérales et aux limonades, elles
        rencontrent toujours plus d'adeptes.</s>
    <s  xml:id="fs2">En effet, notre sondage fait ressortir des ventes nettement
        supérieures à celles de 1987, pour les boissons à
        base de cola notamment.</s>
  </ab>
</div>
```

### 17.5.3 A Three-way Alignment

The preceding encoding of the alignment of parallel passages from two texts requires that those texts and the alignment all be part of the same document. If the texts are in separate documents, then complete URIs, whether absolute or relative (section *17. Linking, Segmentation, and Alignment*), will be required. These external pointers may appear anywhere within the document, but if they are created solely for use in encoding links, they may for convenience be grouped within the <linkGrp> (or other grouping element that uses them for linking).

To demonstrate this facility, we consider how we might encode the alignments in an extract from Comenius' *Orbis Sensualium Pictus*, in the English translation of Charles Hoole (1659). Each topic covered in this work has



(200)

XCVIII.

The Study. *Muſéum.*

The Study 1.
is a place
where a Student, 2.
a part from men,
ſitteth alone,
addicted to his Studies,
whilſt he readeth
Books, 3.
which being within
his reach, he layeth
open upon a Desk 4.
and picketh all the
beſt things out of
them into his own
Manual, 5.

*Muſeum* 1.
eſt locus,
ubi *Studioſus*, 2.
ſecretus ab hominibus,
ſolus ſedet,
*Studiis* deditus,
dum lectitat
*Libros*, 3.
quos penes ſe
ſuper *Pluteum* 4.
exponit, & ex iliis
in *Manuale* 5. ſuum
optima quæq; excerpit,

three parts: a picture, a prose text in Latin describing the topic, and a carefully-aligned translation of the Latin into English, German, or some other vernacular. Key terms in the two texts are typographically distinct, and are linked to the picture by numbers, which appear in the two texts and within the picture as well.

First, we consider the text portions. The English and Latin portions have been encoded as distinct  elements. Identifiers have been attached to each typographic line, but no other encoding added, to simplify the example.

```
<div xml:id="e98" xml:lang="en"
  type="lesson">
  <head>The Study</head>
  <p>
    <seg xml:id="e9801">The Study</seg>
    <seg xml:id="e9802">is a place</seg>
    <seg xml:id="e9803">where a Student,</seg>
    <seg xml:id="e9804">a part from men,</seg>
    <seg xml:id="e9805">sitteth alone,</seg>
    <seg xml:id="e9806">addicted to his Studies,</seg>
    <seg xml:id="e9807">whilst he readeth</seg>
    <seg xml:id="e9808">Books,</seg>
  </p>
</div>
<div xml:id="l98" xml:lang="la"
  type="lesson">
  <head>Muséum</head>
  <p>
    <seg xml:id="l9801">Museum</seg>
    <seg xml:id="l9802">est locus</seg>
    <seg xml:id="l9803">ubi Studiosus,</seg>
    <seg xml:id="l9804">secretus ab hominibus,</seg>
    <seg xml:id="l9805">solus sedet,</seg>
    <seg xml:id="l9806">Studiis deditus,</seg>
    <seg xml:id="l9807">dum lectitat</seg>
    <seg xml:id="l9808">Libros,</seg>
  </p>
</div>
```

Next we consider the non-textual parts of the page. Encoding this requires providing two distinct components: firstly a digitized rendering of the page itself, and secondly a representation of the areas within that image which are to be aligned. In section *12.1. Digital Facsimiles* we present a simple way of doing this using the TEI-defined markup for alignment of facsimiles. In the present chapter we demonstrate a more powerful means of aligning arbitrary polygons and points, which uses the XML notation SVG (see SVG). This provides appropriate facilities for both these requirements:

```
<svg xmlns="http://www.w3.org/2000/svg"
   xmlns:xlink="http://www.w3.org/1999/xlink">
  <image xlink:href="p1764.png" width="597"
    height="897" id="p981"/>
  <rect id="p982" x="75" y="75" width="25"
    height="10"/>
  <rect id="p983" x="55" y="42" width="25"
    height="10"/>
</svg>
```

This example of SVG defines two rectangles at the locations with the specified x and y coordinates. A view is defined on these, enabling them to be mapped by an SVG processor to the image found at the URL specified (`p1764.png`). It also defines unique identifiers for the whole image, and the two views of it, which we will use within our alignment, as shown next (for further discussion of the handling of images and graphics, see section *15.4. Specific Elements for Graphic Images*; for further discussion of using non-TEI XML vocabularies such as SVG within a TEI document, see section *23.8.2. Combining TEI and Non-TEI Modules*).

As printed, the Comenius text exhibits three kinds of alignment.

1 The English and Latin portions are printed in two parallel columns, with corresponding phrases, (represented above by <seg> elements), more or less next to each other.

2 Particular words or phrases are marked as terms in the two languages by a change of rendition: the English text, which otherwise uses black letter type throughout, has the words *The Study*, *a Student*, *Studies*, and *Books* in a roman font; in the Latin text, which is printed in roman, the corresponding words (*Museum*, *Studiosus*, *Studiis*, and *Libros*) are all in italic.

3 Numbered labels appear within the text portions, linking keywords to each other and to sections of the picture. These labels, which have been left out of the above encoding, are attached to the first, third, and last segments in each language quoted below, and also appear (rather indistinctly) within the picture itself. Thus, the images of the study, the student, and his books are each aligned with the correct term for them in the two languages.

The first kind of alignment might be represented by using the *corresp* attribute on the <seg> element. The second kind might be represented by using the <gloss> and <term> mechanism described in section *3.4.1. Terms and Glosses*. The third kind of alignment might be represented using pointers embedded within the texts, for example:

```
<!--... --><seg xml:id="xe9803">where a <ref n="2" target="#xp982">Student</ref>,</seg>
<seg xml:id="xl9803">ubi <ref n="2" target="#xp982">Studiosus</ref>,</seg>
<!--... -->
```

We choose however to use the <link> element, since this provides a more efficient way of representing the three-way alignment between English, Latin, and picture without redundancy.

```
<linkGrp type="alignment">
  <link target="#xe9801 #xl9801 #xp981"/>
  <link target="#xe9802 #xl9802"/>
  <link target="#xe9803 #xl9803 #xp982"/>
  <link target="#xe9804 #xl9804"/>
  <link target="#xe9805 #xl9805"/>
  <link target="#xe9806 #xl9806"/>
  <link target="#xe9807 #xl9807"/>
  <link target="#xe9808 #xl9808 #xp983"/>
</linkGrp>
```

This map, of course, only aligns whole segments and image portions, since these are the only parts of our encoding which bear identifiers and can therefore be pointed to. To add to it the alignment between the typographically distinct words mentioned above, new elements must be defined, either within the text itself or externally by using stand off techniques. Encoding these word pairs as <term> and <gloss>, although intuitively obvious, requires a non-trivial decision as to whether the Latin text is glossing the English, or vice versa. Tagging all the marked words as <term> avoids the difficult decision, but might be thought by some encoders to convey the wrong information about the words in question. Simply tagging them as additional embedded <seg> elements with identifiers that can be aligned like the others is also a possibility.

These solutions all require the addition of further markup to the text. This may pose no problems, or it may be infeasible, for example because the text is held on a read-only medium. If it is not feasible to add more markup to the original text, some form of stand-off markup will be needed. Any item within the text that can be pointed to using the various pointer schemes discussed in this chapter may be used, not simply those which rely on the existence of an *xml:id* attribute. Suppose our example had been more lightly tagged, as follows:

```
<div xml:id="E98" xml:lang="en"
  type="lesson">
  <head>The Study</head>
  <ab>The Study</ab>
```

```
  <ab>is a place</ab>
  <ab>where a Student,</ab>
</div>
<div  xml:id="L98"  xml:lang="la"
  type="lesson">
  <head>Muséum</head>
  <ab>Museum</ab>
  <ab>est locus</ab>
  <ab>ubi Studiosus,</ab>
</div>
```

To express the same alignment mentioned above, we could use an XPath expression to identify the required <ab> elements:

```
<linkGrp  type="alignment">
  <link  target="#xpath(//div[@xml:id='L98']/ab[1]) #xpath(//div[@xml:id='E98']/ab[1])"/>
  <link  target="#xpath(//div[@xml:id='L98']/ab[2]) #xpath(//div[@xml:id='E98']/ab[2])"/>
</linkGrp>
```

In the absence of any markup around individual substrings of the element content, the string-range pointer scheme discussed in *17.2.4.7. string-range()* may also be helpful: for example, to indicate that the words *Studies* and *Studiis* correspond, we might express the link between them as follows:

```
<link  target="#string-range(e9806,16,7) #string-range(l9806,0,7)"/>
```

## 17.6 Identical Elements and Virtual Copies

This section introduces the notion of a *virtual element*, that is, an element which is not explicitly present in a text, but the presence of which an application can infer from the encoding supplied. In this section, we are concerned with virtual elements made by simply cloning existing elements. In the next section (*17.7. Aggregation*), we discuss virtual elements made by aggregating existing elements.

Provided that explicit elements are available to represent the parts or places to be linked, then the global linking attributes *sameAs* and *copyOf* may be used to encode this kind of equivalence:

`att.global.linking` provides a set of attributes for hypertextual linking.

- `@sameAs` points to an element that is the same as the current element.
- `@copyOf` points to an element of which the current element is a copy.

It is useful to be able to represent the fact that one element of text is identical to others, for analytical purposes, or (especially if the elements have lengthy content) to obviate the need to repeat the content. For example, consider the repetition of the <date> element in the following material:

```
<p>In small clumsy letters he wrote:
<q  rend="centered italic">
   <date  xml:id="d840404">April 4th,
       1984</date>.</q>
</p>
<p>He sat back. A sense of complete helplessness had
descended upon him. ...</p>
<p>His small but childish handwriting straggled up
and down the page, shedding first its capital letters
and finally even its full stops:
<q  rend="italic">
```

```
   <date>April 4th, 1984</date>.
    Last night to the flicks. ... </q>
</p>
```

Suppose now that we wish to encode the fact that the second <date> element above has identical content to the first. The *sameAs* attribute is provided for this purpose. Using it, we could recode the last line of the above example as follows:

```
<date  sameAs="#d840404">April 4th,
1984</date>
Last night to the flicks ...
```

The *sameAs* attribute may be used to document the fact that two elements have identical content. It may be regarded as a special kind of link. It should only be attached to an element with identical content to that which it targets, or to one the content of which clearly designates it as a repetition, such as the word *repeat* or *bis* in the representation of the chorus of a song, the second time it is to be sung. The relation specified by the *sameAs* attribute is symmetric: if a chorus is repeated three times and each repetition bears a *sameAs* attribute indicating the first occurrence of the element concerned, it is implied that each chorus is identical, and there is no need for the first occurrence to specify any of its copies.

The *copyOf* attribute is used in a similar way to indicate that the content of the element bearing it is identical to that of another. The difference is that the content is not itself repeated. The effect of this attribute is thus to create a *virtual copy* of the element indicated. Using this attribute, the repeated date in the first example above could be recoded as follows:

```
<date  rend="italic"  copyOf="#d840404"/>
```

An application program should replace whatever is the actual content of an element bearing a *copyOf* attribute with the content of the element specified by it. If the content of the element specified includes other elements, these will become embedded within the element bearing the attribute. Care must be taken to ensure that the document is valid both before and after this embedding takes place. If, for example, the element bearing a *copyOf* attribute requires a mandatory sub-component, then this component must be present (though possibly empty), even though it will be replaced by the content of the targetted element.

The following example demonstrates how the *copyOf* attribute may be used in conjunction with the <seg> element to highlight the differences between almost identical repetitions:

```
<sp>
  <speaker>Mikado</speaker>
  <l>My <seg  xml:id="Mik-L1s">object all sublime</seg>
  </l>
  <l>I shall <seg  xml:id="Mik-L2s">achieve in time</seg>—</l>
  <l  xml:id="Mik-L3">To let <seg  xml:id="L3s">the punishment fit the crime</seg>,</l>
  <l  xml:id="Mik-l4">
    <seg  copyOf="#Mik-L3s"/>;</l>
  <l  xml:id="Mik-l5">And make each pris'ner pent</l>
  <l  xml:id="Mik-l6">Unwillingly represent</l>
  <l  xml:id="Mik-l7">A source <seg  xml:id="Mik-l7s">of innocent merriment</seg>,</l>
  <l  xml:id="Mik-l8">
    <seg  copyOf="#Mik-l7s"/>!</l>
</sp>
<sp>
  <speaker>Chorus</speaker>
```

```
  <l>His <seg  copyOf="#Mik-L1s"/>
  </l>
  <l>He will <seg  copyOf="#Mik-L2s"/>
  </l>
  <l  copyOf="#Mik-L3"/>
  <l  copyOf="#Mik-l4"/>
  <l  copyOf="#Mik-l5"/>
  <l  copyOf="#Mik-l6"/>
  <l  copyOf="#Mik-l7"/>
  <l  copyOf="#Mik-l8"/>
</sp>
```

For further examples of the use of this attribute, see *17.8. Alternation* and *20.3. Another Tree Notation*.

## 17.7 Aggregation

Because of the strict hierarchical organization of elements, or for other reasons, it may not always be possible or desirable to include all the parts of a possibly fragmented text segment within a single element. In section *17.1.4. Intermediate Pointers* we introduced the notion of an intermediate pointer as a way of pointing to discontinuous segments of this kind. In this section we first describe another way of linking the parts of a discontinuous whole, using a set of linking attributes, which are made available for any tag by following the procedure described at the beginning of this chapter. We then describe how the <link> element may be used to aggregate such segments, and finally introduce the <join> element, which is a special-purpose linking element specifically for representing the aggregation of parts, and the <joinGrp> for grouping <join> elements.

The linking attributes for aggregation are *next* and *prev*; each of these attributes has a single identifier as its value:

`att.global.linking` provides a set of attributes for hypertextual linking.

- `@next` (next) points to the next element of a virtual aggregate of which the current element is part.
- `@prev` (previous) points to the previous element of a virtual aggregate of which the current element is part.

It is recommended that the elements indicated by these attributes be of the same type as the element bearing them.

The <join> element is also a member of the class of `att.pointing` elements, and so may carry any of the attributes of that class; for the list, see section *17.1. Links*.

Here is the material on which we base our first illustration of the use of these mechanisms. Our problem is to represent the s-units identified below as qs3 and qs4 as a single (but discontinuous) whole:

```
<q>
  <s  xml:id="qs2">Monsieur Paul, after he has taken equal
     parts of goose breast and the finest pork, and
     broken a certain number of egg yolks into them,
     and ground them <emph>very</emph>, very fine,
     cooks all with seasoning for some three hours.</s>
  <s  xml:id="qs3">
    <emph>But</emph>,</s>
</q>
<s  xml:id="ps2">she pushed her face nearer, and looked with
ferocious gloating at the pâté
inside me, her eyes like X rays,</s>
<q>
  <s  xml:id="qs4">he never stops stirring it!</s>
  <s  xml:id="qs5">Figure to yourself the work of it —</s>
  <s  xml:id="qs6">stir, stir, never stopping!</s>
</q>
```

Using the *prev* and *next* attributes, we can link the s-units with identifiers qs3 and qs4, either singly or doubly as follows:

```
  <s xml:id="qs3" next="#qs4"><emph>But</emph>,</s>
<s xml:id="qs4">he never stops stirring it!</s>
```

```
  <s xml:id="qs3"><emph>But</emph>,</s>
<s xml:id="qs4" prev="#qs3">he never stops stirring it!</s>
```

```
  <s xml:id="qs3" next="#qs4"><emph>But</emph>,</s>
<s xml:id="qs4" prev="#qs3">he never stops stirring it!</s>
```

Double linking of the two s-units, as illustrated by the last of these encodings, is equivalent to specifying a <link> element:

```
<link  type="join"  target="#qs3 #qs4"/>
```

Such a <link> element must carry a *type* attribute with a value of join to specify that the link is to be understood as joining its targets into a single aggregate.

The <join> element is equivalent to a <link> element of type join. Unlike the <link> element, the <join> element can additionally specify information about the virtual element which it represents, by means of its *result* attribute. And finally, unlike the <link> element, the position of a <join> element within a text is significant: it must be supplied at a position where the element indicated by its *result* attribute would be contextually legal.

`<join>` (join) identifies a possibly fragmented segment of text, by pointing at the possibly discontiguous elements which compose it.

`@result` specifies the name of an element which this aggregation may be understood to represent.

`<joinGrp>` (join group) groups a collection of <join> elements and possibly pointers.

`@result` supplies the default value for the *result* on each <join> included within the group.

To conclude the above example, we now use a <join> element to represent the virtual sentence formed by the aggregation of s1 and s2:

```
<join  target="#qs3 #qs4"  result="s"/>
```

As a further example, consider the following list of authors' names. The object of the <join> element here is to provide another list, composed of those authors from the larger list who happen to come from Heidelberg:

```
<list>
  <head>Authors</head>
  <item  xml:id="a_uf">Figge, Udo </item>
  <item  xml:id="a_ch">Heibach, Christiane </item>
  <item  xml:id="a_gh">Heyer, Gerhard </item>
  <item  xml:id="a_bp">Philipp, Bettina </item>
```

```
  <item  xml:id="a_ms">Samiec, Monika </item>
  <item  xml:id="a_ss">Schierholz, Stefan </item>
</list>
<join  target="#a_ch #a_bp #a_ss"
  result="list">
  <desc>Authors from Heidelberg</desc>
</join>
```

The following example shows how <join> can be used to reconstruct a text cited in fragments presented out of order. The poem being remembered (an unusual translation of a well-known poem by Basho) runs 'When the old pond / gets a new frog, / it's a new pond.'

```
<sp>
  <speaker>Hughie</speaker>
  <p>How does it go?
  <q>
     <l  xml:id="frog-x1">da-da-da</l>
     <l  xml:id="frog-L2">gets a new frog</l>
     <l>...</l>
    </q>
  </p>
</sp>
<sp>
  <speaker>Louie</speaker>
  <p>
    <q>
     <l  xml:id="frog-L1">When the old pond</l>
     <l>...</l>
    </q>
  </p>
</sp>
<sp>
  <speaker>Dewey</speaker>
  <p>
    <q>...
    <l  xml:id="frog-L3">It's a new pond.</l>
    </q>
  </p>
  <join  target="#frog-L1 #frog-L2 #frog-L3"
    result="lg"  scope="root"/>
</sp>
```

As with other forms of link, a grouping element <joinGrp> is available for use when a number of <join> elements of the same kind co-occur. This avoids the need to specify the *result* attribute for each <join> if they are all of the same type, and also allows us to restrict the domain within which their target elements are to be found, in the same way as for <linkGrp> elements (see *17.1.3. Groups of Links*). Like a <join>, a <joinGrp> may appear only where the elements represented by its contents are legal. Thus if we had created many <join> tags of the sort just described, we could group them together, and require that their components are all contained by an element with the identifier MFKFhungry as follows:

```
<joinGrp  domains="#mfkfhungry #mfkfhungry"
  result="s">
  <join  target="#qs3 #qs4"/>
```

```
  <join  target="#qs5 #qs6"/>
</joinGrp>
```

The <join> element is useful as a means of representing non-hierarchic structures (as further discussed in chapter *21. Non-hierarchical Structures*). It may also be used as a convenient way of representing a variety of analytic units, like the  and <interp> elements discussed in chapter *18. Simple Analytic Mechanisms*. As an example, consider the following famous Zen koan:

> Zui-Gan called out to himself every day, 'Master.'
>
> Then he answered himself, 'Yes, sir.'
>
> And then he added, 'Become sober.'
>
> Again he answered, 'Yes, sir.'
>
> 'And after that,' he continued, 'do not be deceived by others.'
>
> 'Yes, sir; yes, sir,' he replied.

Suppose now that we wish to represent an interpretation of the above passage in which we distinguish between the various 'voices' adopted by Zui-Gan. In the following encoding, the *who* attribute has been used for this purpose; its value on each occasion supplies a pointer to the 'voice' to which each speech is attributed. (For convenience in this example, we use simply the first occurrence of the names used for each voice as the target for these pointers.) Note also that we add *xml:id* attributes to each distinct speech fragment, which we can then use to link the material spoken by each voice:

```
<text  xml:id="zuitxt">
  <body>
    <p>
      <name  xml:id="zuigan">Zui-Gan</name> called out to himself every day,
    <q  next="#zuiq2"  xml:id="zuiq1"
       who="#zuigan">
       <name  xml:id="master">Master</name>.</q>
    </p>
    <p>Then he answered himself,
    <q  next="#zuiq4"  xml:id="zuiq2"
       who="#zuigan">Yes, sir.</q>
    </p>
    <p>And then he added,
    <q  next="#zuiq5"  xml:id="zuiq3"
       who="#master">Become sober.</q>
    </p>
    <p>Again he answered,
    <q  next="#zuiq7"  xml:id="zuiq4"
       who="#zuigan">Yes, sir.</q>
    </p>
    <p>
      <q  next="#zuiq6"  xml:id="zuiq5"
        who="#master">And after that,</q>
         he continued,
    <q  xml:id="zuiq6"  who="#master">do not be deceived by others.</q>
    </p>
    <p>
      <q  xml:id="zuiq7"  who="#zuigan">Yes, sir; yes, sir,</q>
         he replied.</p>
  </body>
</text>
```

However, by using the <join> element, we can directly represent the complete speech attributed to each voice:

```
<joinGrp  result="q">
  <join  target="#zuiq1 #zuiq2 #zuiq4 #zuiq7">
    <desc>what Zui-Gan said</desc>
  </join>
  <join  target="#zuiq3 #zuiq5 #zuiq6">
    <desc>what Master said</desc>
  </join>
</joinGrp>
```

Note the use of the <desc> child element within the two <join>s making up the <q> element here. These enable us to document the speakers of the two virtual <q> elements represented by the <join> elements; this is necessary because the there is no way of specifying the attributes to be associated with a virtual element, in particular there is no way to specify a *who* value for them.

Suppose now that *xml:id* attributes, for whatever reasons, are not available. Then <ptr> elements may be created using any of the methods described in section *17.2.4. TEI XPointer Schemes.* The *xml:id* attributes of *these* elements may now be specified by the *target* attribute on the <join> elements.

```
<text>
  <body>
<!-- five div1 elements -->
    <div1>
      <p>Zui-Gan called out to himself every day, <q>Master.</q>
      </p>
      <p>Then he answered himself, <q>Yes, sir.</q>
      </p>
      <p>And then he added, <q>Become sober.</q>
      </p>
      <p>Again he answered, <q>Yes, sir.</q>
      </p>
      <p>
        <q>And after that,</q> he continued, <q>do not be deceived by others.</q>
      </p>
      <p>
        <q>Yes, sir; yes, sir,</q> he replied.</p>
      <ab  type="aggregation">
        <ptr  xml:id="rzuiq1"
          target="#xpath(//div1[6]/p[1]/q[1])"/>
        <ptr  xml:id="rzuiq2"
          target="#xpath(//div1[6]/p[2]/q[1])"/>
        <ptr  xml:id="rzuiq3"
          target="#xpath(//div1[6]/p[3]/q[1])"/>
        <ptr  xml:id="rzuiq4"
          target="#xpath(//div1[6]/p[4]/q[1])"/>
        <ptr  xml:id="rzuiq5"
          target="#xpath(//div1[6]/p[5]/q[1])"/>
        <ptr  xml:id="rzuiq6"
          target="#xpath(//div1[6]/p[5]/q[2])"/>
        <ptr  xml:id="rzuiq7"
          target="#xpath(//div1[6]/p[6]/q[1])"/>
        <joinGrp  evaluate="one"  result="q">
          <join  target="#rzuiq1 #rzuiq2 #rzuiq4 #rzuiq7">
            <desc>what Zui-Gan said</desc>
          </join>
          <join  target="#rzuiq3 #rzuiq5 #rzuiq6">
```

```
          <desc>what Master said</desc>
        </join>
      </joinGrp>
    </ab>
  </div1>
 </body>
</text>
```

The extended pointer with identifier rzuiq2, for example, may be read as 'the first <q> in the first <p>, within the sixth <div1> element of the current document.'

## 17.8 Alternation

This section proposes elements for the representation of alternation. We say that two or more elements are in *exclusive alternation* if any of those elements could be present in a text, but one and only one of them is; in addition, we say that those elements are *mutually exclusive*. We say that the elements are in *inclusive alternation* if at least one (and possibly more) of them is present. The elements that are in alternation may also be called *alternants*.

The need to mark exclusive alternation arises frequently in text encoding. A common situation is one in which it can be determined that exactly one of several different words appears in a given location, but it cannot be determined which one. One way to mark such an exclusive alternation is to use the linking attribute *exclude*. Having marked an exclusive alternation, it can sometimes later be determined which of the alternants actually appears in the given location. To preserve the fact that an alternation was posited, one can add the linking attribute *select* to a tag which hierarchically encompasses the alternants, which points to the one which actually appears. To assign responsibility and degree of certainty to the choice, one can use the <certainty> tag described in chapter *22. Certainty, Precision, and Responsibility*. Also see that chapter for further discussion of certainty in general.

The *exclude* and *select* attributes may be used with any element assuming that they have been declared following the procedure discussed in the introduction to this chapter.

`att.global.linking` provides a set of attributes for hypertextual linking.

- `@exclude` points to elements that are in exclusive alternation with the current element.
- `@select` selects one or more alternants; if one alternant is selected, the ambiguity or uncertainty is marked as resolved. If more than one alternant is selected, the degree of ambiguity or uncertainty is marked as reduced by the number of alternants not selected.

A more general way to mark alternation, encompassing both exclusive and inclusive alternation, is to use the linking element <alt>. The description and attributes of this tag and of the associated grouping tag <altGrp> are as follows. These elements are also members of the `att.pointing` class and therefore have all the attributes associated with that class.

`<alt>` (alternation) identifies an alternation or a set of choices among elements or passages.

- `@weights` If *mode* is excl, each weight states the probability that the corresponding alternative occurs. If *mode* is incl each weight states the probability that the corresponding alternative occurs given that at least one of the other alternatives occurs.

`<altGrp>` (alternation group) groups a collection of <alt> elements and possibly pointers.

To take a simple hypothetical example, suppose in transcribing a spoken text, we encounter an utterance that we can understand either as *We had fun at the beach today.* or as *We had sun at the beach today.* We can represent the exclusive alternation of these two possibilities by means of the *exclude* attribute as follows.

```
<div type="interview">
  <u exclude="#we.sun1" xml:id="we.fun1">We had fun at the beach today.</u>
  <u exclude="#we.fun1" xml:id="we.sun1">We had sun at the beach today.</u>
</div>
```

If it is then determined that the speaker said *fun*, not *sun*, the encoder could amend the text by deleting the alternant containing *sun* and the *exclude* attribute on the remaining alternant. Alternatively, the encoder could preserve the

fact that there was uncertainty in the original transcription by retaining the alternants, and assigning the we.fun value to the *select* attribute value on the  element that encompasses the alternants, as in:

```
<div  select="#we.fun2"  type="interview">
  <u  exclude="#we.sun2"  xml:id="we.fun2">We had fun at the beach
     today.</u>
  <u  exclude="#we.fun2"  xml:id="we.sun2">We had sun at the beach today.</u>
</div>
```

The above alternation (including the *select* attribute) could be recoded by assigning the *exclude* attributes to tags that enclose just the words or even the characters that are mutually exclusive, as in:[103]

```
<div  type="interview">
  <u  select="#fun3">We had
  <seg  exclude="#sun3"  xml:id="fun3"
     type="word">fun</seg>
    <seg  exclude="#fun3"  xml:id="sun3"
     type="word">sun</seg>
    at the beach today.</u>
</div>
```

```
<div  type="interview">
  <u>We had
  <seg  select="#id-f"  type="word">
     <seg  exclude="#id-s"  xml:id="id-f"
       type="character">f</seg>
     <seg  exclude="#id-f"  xml:id="id-s"
       type="character">s</seg>
        un</seg>
    at the beach today.</u>
</div>
```

Now suppose that the transcriber is uncertain whether the first word in the utterance is *We* or *Lee*, but is certain that if it is *Lee*, then the other uncertain word is definitely *fun* and not *sun.* The three utterances that are in mutual exclusion can be encoded as follows.

```
<div  type="interview">
<!-- ... -->
  <u  exclude="#we.sun4 #lee.fun4"
    xml:id="we.fun4">We had fun at the beach today.</u>
  <u  exclude="#we.fun4 #lee.fun4"
    xml:id="we.sun4">We had sun at the beach today.</u>
  <u  exclude="#we.fun4 #we.sun4"
    xml:id="lee.fun4">Lee had fun at the beach today.</u>
<!-- ... -->
</div>
```

The preceding example can also be encoded with *exclude* attributes on the word segments *We*, *Lee*, *fun*, and *sun*:

[103] See section *18.1. Linguistic Segment Categories* for discussion of the <w> and <c> tags that can be used in the following examples instead of the <seg type="word"> and <seg type="character"> tags.

```
<u>
  <seg  exclude="#lee"  xml:id="we"  type="word">We</seg>
  <seg  exclude="#we #sun"  xml:id="lee"
    type="word">Lee</seg>
had
<seg  exclude="#sun"  xml:id="fun"
   type="word">fun</seg>
  <seg  exclude="#fun #lee"  xml:id="sun"
    type="word">sun</seg>
at the beach today.
</u>
```

The value of the *select* attribute is defined as a list of identifiers; hence it can also be used to narrow down the range of alternants, as in:

```
<div  select="#we.fun5 #lee.fun5"
  type="interview">
  <u  exclude="#we.sun5 #lee.fun5"
    xml:id="we.fun5">We had fun at the beach today.</u>
  <u  exclude="#we.fun5 #lee.fun5"
    xml:id="we.sun5">We had sun at the beach today.</u>
  <u  exclude="#we.fun5 #we.sun5"
    xml:id="lee.fun5">Lee had fun at the beach today.</u>
</div>
```

This is interpreted to mean that either the first or the third <u> element tag appears, and is thus equivalent to just the alternation of those two tags:

```
<div  type="interview">
  <u  exclude="#lee.fun6"  xml:id="we.fun6">We had fun at the beach
     today.</u>
  <u  exclude="#we.fun6"  xml:id="lee.fun6">Lee had fun at the beach today.</u>
</div>
```

The *exclude* attribute can also be used in case there is uncertainty about the tag that appears in a certain position. For example, the occurrence of the word *May* in the s-unit *Let's go to May* can be interpreted, in the absence of other information, either as a person's name or as a date. The uncertainty can be rendered as follows, using the *exclude* attribute.

```
<s>Let's go to
<name  exclude="#mayn"  xml:id="mayd">May</name>
  <date  copyOf="#mayd"  exclude="#mayd"
    xml:id="mayn"/>.</s>
```

Note the use of the *copyOf* attribute discussed in section *17.6. Identical Elements and Virtual Copies*; this avoids having to repeat the content of the element whose correct tagging is in doubt.

The *copyOf* and the *exclude* attributes also provide for a simple way of indicating uncertainty about exactly where a particular element occurs in a document.[104] For example suppose that a particular <div2> element appears either as the third and last of the <div2> elements within the first <div1> element in the body of a document, or as the first <div2> of the second <div1>. One solution would be to record the <div2> in its entirety in the first of these positions, and a virtual copy of it in the second, and mark them as excluding each other as follows:

[104] An alternative way of representing this problem is discussed in chapter *22. Certainty, Precision, and Responsibility.*

```
<body>
  <div1 xml:id="C1">
    <div2 xml:id="C1S3" exclude="#C2S1"/>
  </div1>
  <div1 xml:id="C2">
    <div2 xml:id="C2S1" copyOf="#C1S3"
      exclude="#C1S3"/>
  </div1>
</body>
```

In this case, the *select* attribute, if used, would appear on the <body> element.

Mutual exclusion can also be expressed using a <link>; the first example in this section can be recoded by removing the *exclude* attributes from the <u> elements, and adding a <link> element as follows:[105]

```
<div type="interview">
  <u xml:id="we.had.fun">We had fun at the beach today.</u>
  <u xml:id="we.had.sun">We had sun at the beach today.</u>
  <link type="exclusiveAlternation"
    target="#we.had.fun #we.had.sun"/>
</div>
```

Now we define the specialized linking element <alt>, making it a member of the class `att.pointing`, and assigning it a *mode* attribute, which can have either of the values excl (for exclusive) or incl (for inclusive). Then the following equivalence holds:

```
<alt target="#a #b" mode="excl"/>
```

=

```
<link target="#a #b"
  type="exclusive_alternation"/>
```

The preceding <link> element may therefore be recoded as the following <alt> element.

```
<alt target="#we.had.fun #we.had.sun"
  mode="excl"/>
```

Another attribute that is defined specifically for the <alt> element is *weights*, which is to be used if one wishes to assign *probabilistic weights* to the targets (alternants). Its value is a list of numbers, corresponding to the targets, expressing the probability that each target appears. If the alternants are mutually exclusive, then the weights must sum to 1.0.

Suppose in the preceding example that it is equiprobable whether *fun* or *sun* appears. Then the <alt> element that represents the alternation may be stated as follows:

```
<alt target="#we.fun #we.had.sun"
  mode="excl" weights="0.5 0.5"/>
```

[105] In this example, we have placed the <link> next to the elements that represent the alternants. It could also have been placed elsewhere in the document, perhaps within a <linkGrp>.

The assignment of a weight of 1.0 to one target (and weights of 0 to all the others) is equivalent to selecting that target. Thus the following encoding is equivalent to the second example at the beginning of this section.

```
<u  xml:id="we.fun">We had fun at the beach today.</u>
<u  xml:id="we.sun">We had sun at the beach today.</u>
<alt  target="#we.fun #we.sun"  mode="excl"
  weights="1 0"/>
```

The sum of the weights for <alt mode="incl"> tags ranges from 0% to (100 × k)%, where k is the number of targets. If the sum is 0%, then the alternation is equivalent to exclusive alternation; if the sum is (100 x k)%, then all of the alternants must appear, and the situation is better encoded without an <alt> tag.

If it is desired, <alt> elements may be grouped together in an <altGrp> element, and attribute values shared by the individual <alt> elements may be identified on the <altGrp> element. The *targFunc* attribute defaults to the value first.alternant next.alternant.

To illustrate, consider again the example of a transcribed utterance, in which it is uncertain whether the first word is *We* or *Lee*, whether the third word is *fun* or *sun*, but that if the first word is *Lee*, then the third word is *fun*. Now suppose we have the following additional information: if *we* occurs, then the probability that *fun* occurs is 50% and that *sun* occurs is 50%; if *fun* occurs, then the probability that *we* occurs is 40% and that *Lee* occurs is 60%. This situation can be encoded as follows.

```
<u>
  <seg  exclude="#lee2"  xml:id="we2"
    type="word">We</seg>
  <seg  exclude="#we2"  xml:id="lee2"
    type="word">Lee</seg>
had
<seg  exclude="#sun2"  xml:id="fun2"
    type="word">fun</seg>
  <seg  exclude="#fun2"  xml:id="sun2"
    type="word">sun</seg>
at the beach today.
</u>
<altGrp>
  <alt  target="#we2 #lee2"/>
  <alt  target="#fun2 #sun2"/>
  <alt  target="#we2 #fun2"  mode="incl"
    weights="0.5 0.5"/>
  <alt  target="#lee2 #fun2"  mode="incl"
    weights="1.0 0.6"/>
</altGrp>
```

As noted above, when the *mode* attribute on an <alt> has the value incl, then each weight states the probability that the corresponding alternative occurs, given that at least one of the other alternatives occurs.

From the information in this encoding, we can determine that the probability is about 28.5% that the utterance is 'We had fun at the beach today', 28.5% that it is *We had sun at the beach today*, and 43% that it is *Lee had fun at the beach today*.

Another very similar example is the following regarding the text of a Broadway song. In three different versions of the song, the same line reads 'Her skin is tender as a leather glove', 'Her skin is tender as a baseball glove', and 'Her skin is tender as Dimaggio's glove.'[106]

If we wish to express this textual variation using the <alt> element, we can record our relative confidence in the readings *Dimaggio's* (with probability 50%), *a leather* (25%), and *a baseball* (25%).

Let us extend the example with a further (imaginary) variation, supposing for the sake of the argument that the next line is variously given as *and she bats from right to left* (with probability 50%) or *now ain't that too damn bad*

[106]The variant readings are found in the commercial sheet music, the performance score, and the Broadway cast recording.

(with probability 50%). Using the <alt> element, we can express the conviction that if the first choice for the second line is correct, then the probability that the first line contains *Dimaggio's* is 90%, and each of the others 5%; whereas if the second choice for the second line is correct, then the probability that the first line contains *Dimaggio's* is 10%, and each of the others is 45%. This can be encoded, with an <altGrp> element containing a combination of exclusive and inclusive <alt> elements, as follows.

```
<div xml:id="bm" type="song">
  <l>Her skin is tender as
  <seg xml:id="dm">Dimaggio's</seg>
    <seg xml:id="lt">a leather</seg>
    <seg xml:id="bb">a baseball</seg>
     glove,</l>
  <l xml:id="rl">and she bats from right to left.</l>
  <l xml:id="db">now ain't that too damn bad.</l>
</div>
<altGrp>
  <alt target="#dm #lt #bb" mode="excl"
    weights="0.5 0.25 0.25"/>
  <alt target="#rl #db" mode="excl"
    weights="0.50 0.50"/>
</altGrp>
<altGrp mode="incl">
  <alt target="#dm #rl" weights="0.90 0.90"/>
  <alt target="#lt #rl" weights="0.5 0.5"/>
  <alt target="#bb #rl" weights="0.5 0.5"/>
  <alt target="#dm #db" weights="0.10 0.10"/>
  <alt target="#lt #db" weights="0.45 0.90"/>
  <alt target="#bb #db" weights="0.45 0.90"/>
</altGrp>
```

## 17.9 Stand-off Markup

### 17.9.1 Introduction

Most of the mechanisms defined in this chapter rely to a greater or lesser extent on the fact that tags in a marked-up document can both assert a property for a span of text which they enclose, and assert the existence of an association between themselves and some other span of text elsewhere. In stand-off markup, there is a clear separation of these two behaviours: the markup does not directly contain any part of the text, but instead includes it by reference. One specific mechanism recommended by these Guidelines for this purpose is the standard XInclude mechanism defined by the W3C; another is to use pointers as demonstrated elsewhere in this chapter.

There are many reasons for using stand-off markup: the source text might be read-only so that additional markup cannot be added, or a single text may need to be marked up according to several hierarchically incompatible schemes, or a single scheme may need to accommodate multiple hierarchical ambiguities, so that a single markup tree is not the most faithful representation of the source material.

This section describes a generic mechanism for expressing *all* kinds of markup externally as stand-off tags, for use whenever it is appropriate; and a place in the TEI structure (<standOff>) to contain certain common kinds of stand-off markup.

Throughout this section the following terms will be systematically used in specific senses.

***source document*** a document to which the stand-off markup refers (a source document can be either XML or plain text); there may be more than one source document.

***internal markup*** markup that is already present in an XML source document

***stand-off markup*** markup that is either outside of the source document and points in to it to the data it describes, or is pointed at by the data that refers to it; or alternatively is in another part of the source document and points elsewhere within the document to the data it describes, or is pointed at by data elsewhere that refers to it.

***external document*** a document that contains stand-off markup that points to a different, source document

***internalize*** the action of creating a new XML document with external markup and data integrated with the source document data, and possibly some source document markup as well

***externalize*** a process applied to markup from a pre-existing XML document, which splits it into two documents, an XML (external) document containing some of the markup of the original document, and another (source) XML document containing whatever text content and markup has not been extracted into the stand-off document; if all markup has been externalized from a document, the new source may be a plain text document

The three major requirements satisfied by this scheme for stand-off markup are:

a any valid TEI markup can be either internal or external,

b external markup can be internalized by applying it to the document content by either substituting the existing markup or adding to it, to form a valid TEI document, and

c the external markup itself specifies whether an internalized document is to be created by substituting the existing internal markup or by adding to it.

### 17.9.2 Overview of XInclude

Stand-off markup which relies on the inclusion of virtual content is adequately supported by the W3C XInclude recommendation, which is also recommended for use by these Guidelines.[107] XInclude defines a namespace (*http://www.w3.org/2001/XInclude*), which in these Guidelines will be associated with the prefix *xi:*, and exactly two elements, `<xi:include>` and `<xi:fallback>`. XInclude relies on the XPointer framework discussed elsewhere in this chapter to point to the actual fragments of text to be internalized. Although XInclude only requires support for the `element()` scheme of XPointer, these Guidelines permit the use of any of the pointing schemes discussed in section *17.2. Pointing Mechanisms*.

XInclude is a W3C recommendation which specifies a syntax for the inclusion within an XML document of data fragments placed in different resources. Included resources can be either plain text or XML. XInclude instructions within an XML document are meant to be replaced by a resource targetted by a URI, possibly augmented by an XPointer that identifies the exact subresource to be included.

The `<xi:include>` element uses the *href* attribute to specify the location of the resource to be included; its value is an URI containing, if necessary, an XPointer. Additionally, it uses the *parse* attribute (whose only valid values are text and xml) to specify whether the included content is plain text or an XML fragment, and the *encoding* attribute to provide a hint, when the included fragment is text, of the character encoding of the fragment. An optional `<xi:fallback>` element is also permitted within an `<xi:include>`; it specifies alternative content to be used when the external resource cannot be fetched for some reason. Its use is not however recommended for stand-off markup.

### 17.9.3 Stand-off Markup in TEI

The operations of internalizing and externalizing markup are very useful and practically important. XInclude processing as defined by the W3C *is* internalization of one or more source documents' content into a stand-off document. TEI use of XInclude for stand-off markup enables use of XInclude-conformant software to perform this useful operation. However, internalization is not clearly defined for all stand-off files, because the structure of the internal and external markup trees may overlap. In particular, when an external markup document selects a range that overlaps partial elements in the source document, it is not clear how the semantics of internalization (inclusion) should work, since partial elements are not XML objects.[108] XInclude defines a semantics for this case that involves only complete elements.

When a range selection partially overlaps a number of elements in a source document, XInclude specifies that the partially overlapping elements should be included as well as all completely overlapping elements and characters (partially overlapping characters are not possible). The effect of this is that elements that straddle the start or end of a selected range will be included as wrappers for those of their children that are completely or partially selected by the range. For example, given the following source document:

[107] The version on which this text is based is the W3C Recommendation dated 20 December 2004..

[108] This corresponds to the observation that overlapping XML tags reflecting a textual version of such an inclusion would not even be well-formed XML. This kind of overlap in textual phenomena of interest is in fact the major reason that stand-off markup is needed.

```
<body>
  <p  xml:id="par1">home, <emph>home</emph> on Brokeback Mountain.</p>
  <p  xml:id="par2">That was the <emph>song</emph> that I sang</p>
</body>
```

and the following external document:

```
  <body>
<div><include href="example1.xml" xmlns="http://www.w3.org/2001/XInclude"
xpointer="range(xpath(id('par1')//emph),xpath(id('par2')//emph))"/>
</div>
</body>
```

the resulting document after XInclude processing of this external document would be:

```
  <body>
<div>
<p xml:id="par1">home, <emph>home</emph> on Brokeback Mountain.</p>
<p xml:id="par2">That was the <emph>song</emph> that I sang</p>
</div>
</body>
```

The result of the inclusion is two paragraph elements, while the original range designated in the source document overlapped two paragraph fragments. The semantics of XInclude require the creation of well-formed XML results even though the pointing mechanisms it uses do not necessarily respect the hierarchical structure of XML documents, as in this case. While this is a good way to ensure that internalization is always possible, it has implications for the use of XInclude as a notation for the *description* of overlapping markup structures.

When overlapping hierarchies need to be represented for a single document, each hierarchy must be represented by a separate set of XInclude tags pointing to a common source document. This sort of structure corresponds to common practice in work with linguistic text corpora. In such corpora, each potentially overlapping hierarchy of elements for the text is represented as a separate stream of stand-off markup. Generally the source text contains markup for the smallest significant units of analysis in the corpus, such as words or morphemes, this information and its markup representing a layer of common information that is shared by all the various hierarchies. As a way of organizing the representation of complex data, this technique generally allows a large number of *xml:id* attributes to be attached to the shared elements, providing robust anchors for links and facilitating adjustments to the source document without breaking external documents that reference it.

Any tag can be externalized by removing its content and replacing it with an `<xi:include>` element that contains an XPointer pointing to the desired content.

For instance the following portion of a TEI document:

```
<text>
  <body>
    <head>1755</head>
    <l>To make a prairie it takes a clover and one bee,</l>
    <l>One clover, and a bee,</l>
    <l>And revery.</l>
    <l>The revery alone will do,</l>
    <l>If bees are few.</l>
  </body>
</text>
```

can be externalized by placing the actual text in a separate document, and providing exactly the same markup with the <xi:include> elements:
**Source.xml**

```
  <content>To make a prairie it takes a clover and one bee,\n
One clover, and a bee,\n
And revery.\n
The revery alone will do,\n
If bees are few.\n
</content>
```

**External.xml**

```
  <text xmlns:xi="http://www.w3.org/2001/XInclude">
<body>
<head>1755</head>
<l>
<xi:include href="Source.xml" parse="xml"
xpointer="string-range(element(/1),  0, 48)"/>
</l>
<l>
<xi:include href="Source.xml" parse="xml"
xpointer="string-range(element(/1), 49, 71)"/>
</l>
<l>
<xi:include href="Source.xml" parse="xml"
xpointer="string-range(element(/1), 72, 83)"/>
</l>
<l>
<xi:include href="Source.xml" parse="xml"
xpointer="string-range(element(/1), 84,109)"/>
</l>
<l>
<xi:include href="Source.xml" parse="xml"
xpointer="string-range(element(/1),110,126)"/>
</l>
</body>
</text>
```

Please note that this specification requires that the XInclude namespace declaration is present in all cases. The <xi:fallback> element contains text or XML fragments to be placed in the document if the inclusion fails for any reason (for instance due to inaccessibility of an external resource). The <xi:fallback> element is optional; if it is not present an XInclude processor must signal a fatal error when a resource is not found. This is the preferred behaviour for use with stand-off markup. These Guidelines recommend against the use of <xi:fallback> for stand-off markup.

### 17.9.4 Well-formedness and Validity of Stand-off Markup

The whole source fragment identified by an XInclude element, as well as any markup therein contained is inserted in the position specified, and an XInclude processor is required to ensure that the resulting internalized document is well-formed. This has obvious implications when the external document contains XML markup. A plain text source document will always create a well-formed internalized document.

While a TEI customization may permit `<xi:include>` elements in various places in a TEI document instance, in general these Guidelines suggest that validity be verified after the resolution of all the `<xi:include>` elements.

### 17.9.5 Including Text or XML Fragments

When the source text is plain text the overall form of the XPointer pointing to it is of minimal importance. The form of the XPointer matters considerably, on the other hand, when the source document is XML.

In this case, it is rather important to distinguish whether we intend to substitute the source XML with the new one, or just to add new markup to it. The XPointers used in the references can express both cases.

A simple way is to make sure to select only textual data in the XPointer. For instance, given the following document:

**Source.xhtml**

```
<xhtml:html>
  <xhtml:body>
    <xhtml:div>To make a prairie it takes a <xhtml:a  href="clover.gif">clover</xhtml:a>
        and one <xhtml:a  href="bee.gif">bee</xhtml:a>,</xhtml:div>
    <xhtml:div>One <xhtml:a  href="clover.gif">clover</xhtml:a>, and
        a <xhtml:a  href="bee.gif">bee</xhtml:a>,</xhtml:div>
    <xhtml:div>And revery.</xhtml:div>
    <xhtml:div>The revery alone will do,</xhtml:div>
    <xhtml:div>If bees are few.</xhtml:div>
  </xhtml:body>
</xhtml:html>
```

the expression `range(element(/1/2/1.0),element(/1/2/11.1))` will select the whole poem, text content *and*  elements *and* hypertext links (NB: in XPointer whitespace-only text nodes count).

On the contrary, the expressions `xpointer(//text()/range-to(.))` and `xpointer(string-range(//text()`, will only select the text of the poem, with no markup inside.

Thus, the following could be a valid stand-off document for the *Source.xhtml* document:

**External2.xml**

```
  <text xmlns:xi="http://www.w3.org/2001/XInclude">
<body>
<head>1755</head>
<l>
<xi:include href="Source.xhtml"
xpointer='xpointer(string-range(//div[1]/text(),"To")/range-to(//div[1]/text(),"bee,")'/>
</l>
<l>
<xi:include href="Source.xhtml"
xpointer='xpointer(string-range(//div[2]/text(),"One")/range-to(//div[2]/text(),"bee,")'/>
</l>
<l>
<xi:include href="Source.xhtml"
xpointer='xpointer(string-range(//div[3]/text(),"And")/range-to(//div[3]/text(),".")'/>
</l>
<l>
<xi:include href="Source.xhtml"
xpointer='xpointer(string-range(//div[4]/text(),"The")/range-to(//div[4]/text(),",")'/>
</l>
<l>
<xi:include href="Source.xhtml"
xpointer='xpointer(string-range(//div[5]/text(),"If")/range-to(//div[5]/text(),".")'/>
</l>
```

```
</body>
</text>
```

## 17.10 The <standOff> Container

The <standOff> element is intended to hold content that does not fit well in the <text> (e.g. because it is not transcribed from the source), nor in the <teiHeader> (e.g. because it is not metadata about the source or transcription). Examples include contextual information about named entities (typically encoded using <listBibl>, <listOrg>, <listNym>, <listPerson>, or <listPlace>), annotations indicating the morphosyntactic features of a text, and annotations commenting on or associating parts of a text with additional information.

`<standOff>` Functions as a container element for linked data, contextual information, and stand-off annotations embedded in a TEI document.

As a member of `model.resource`, <standOff> may occur as a child of <TEI> (or <teiCorpus>). If the metadata that describes the <standOff> is largely the same as the metadata that describes the associated resource (e.g., the transcribed text in <text>), then the <standOff> and the encoded associated resource may appear as children of the same <TEI> element. The example below has a transcription with <placeName> elements in the text linked to a list of <place> elements in the <standOff> section.

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!-- ... -->
  </teiHeader>
  <standOff>
    <listPlace>
<!-- ... -->
     <place  xml:id="Cilicia">
       <placeName>Cilicia</placeName>
       <idno  type="URI">https://pleiades.stoa.org/places/658440</idno>
     </place>
     <place  xml:id="Creta">
       <placeName  xml:lang="la">Creta</placeName>
       <placeName  xml:lang="en">Crete</placeName>
       <idno  type="URI">https://pleiades.stoa.org/places/589748</idno>
     </place>
<!-- ... -->
     <place  xml:id="Rhodus">
       <placeName  xml:lang="la">Rhodus</placeName>
       <placeName  xml:lang="en">Rhodes</placeName>
       <idno  type="URI">https://pleiades.stoa.org/places/590031</idno>
     </place>
     <place  xml:id="Syria">
       <placeName>Syria</placeName>
       <idno  type="URI">https://pleiades.stoa.org/places/1306</idno>
     </place>
<!-- ... -->
    </listPlace>
  </standOff>
  <text>
    <body>
      <div  type="edition"
        xml:id="edition-text">
        <div  type="textpart"  n="1"
          xml:id="part1">
          <head>Bellum Alexandrinum</head>
          <p  n="1"  xml:id="p1">
```

```
          <seg  n="1"  xml:id="seg-1.1">Bello Alexandrino conflato
<persName  ref="#Caesar">Caesar</persName>
            <app>
              <lem>
                <placeName  ref="#Rhodus">Rhodo</placeName>
              </lem>
              <rdg  wit="#S"  ana="#orthographical">Ordo</rdg>
            </app> atque ex <placeName  ref="#Syria">Syria</placeName>
            <placeName  ref="#Cilicia">Cilicia</placeName>que omnem classem arcessit;
<app>
              <lem>
                <placeName  ref="#Creta">Creta</placeName>
              </lem>
              <rdg  wit="#S"  ana="#orthographical">certa</rdg>
            </app> sagittarios, equites ab rege <orgName  ref="#Nabataei">Nabataeorum</orgName>
            <persName  ref="#Malchus">Malcho</persName> euocat; tormenta undique
conquiri et
                     frumentum mitti, auxilia adduci iubet.</seg>
<!-- ... -->
        </p>
<!-- ... -->
       </div>
<!-- ... -->
     </div>
   </body>
  </text>
</TEI>
```

## 17.11 Annotations

The <annotation> element's structure is based on the Web Annotation Data Model (WADM). A Web Annotation may have one or more targets, which are URIs, and zero or more bodies, which may be either URIs or embedded text. A Web Annotation may also contain metadata about the annotation, such as the creator, creation and modification dates, and license information. The <annotation> element implements a subset of WADM, using TEI elements and attributes to encode the same information, with a focus on annotating TEI documents. Targets are represented using the *target* attribute on <annotation>. URI bodies are represented using <ref> or <ptr> and embedded text bodies using <note>. Lifecycle and license information may be given using <respStmt>, <revisionDesc>, and <licence>.

`<listAnnotation>` contains a list of annotations, typically encoded as <annotation>, <annotationBlock>, or <note>, possibly organized with nested <listAnnotation> elements.

`<annotation>` represents an annotation following the Web Annotation Data Model.

TEI annotations are, in general, intended to capture the output of processes that annotate TEI texts without altering the text and markup in the <text>. They allow this kind of output to be represented directly in TEI, and thus to be processed using the same toolchains as the texts they annotate. A named entity recognition workflow might use <annotation>, for example to associate names in the text with <person> elements instead of attempting to rewrite the TEI text using inline <persName>. Projects may wish to use this mechanism to layer information onto a TEI text in case where using inline elements might result in complicated markup.

The example below illustrates how stand-off annotations can be used to connect words in a text with <place> elements in a list. The words 'Gallia' and 'Galliae' in the edition are connected by an annotation in the <standOff> section which points to them using `string-range()` pointers and references the definition of the place (also in the <standOff> section). If the set of annotations were created in a process separate from the creation of the transcription and then imported into the transcription document, then wrapping them in a <TEI> element with its own <teiHeader> providing metadata for the annotations might be advisable.

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>De Bello Gallico</title>
      </titleStmt>
<!-- ... -->
    </fileDesc>
  </teiHeader>
  <text>
    <body>
      <div type="edition">
        <div type="textpart" subtype="chapter"
          n="1" xml:id="ch1">
          <p n="1" xml:id="c1p1">
            <seg n="1" xml:id="c1p1s1">Gallia est omnis divisa in partes tres, quarum
unam
                      incolunt Belgae, aliam Aquitani, tertiam qui ipsorum lingua
Celtae, nostra
                      Galli appellantur.</seg>
<!-- ... -->
            <seg n="6" xml:id="c1p1s6">Belgae ab extremis Galliae finibus oriuntur,
                      pertinent ad inferiorem partem fluminis Rheni, spectant in
septentrionem et
                      orientem solem.</seg>
            <seg n="7" xml:id="c1p1s7">Aquitania a Garumna flumine ad Pyrenaeos montes
et
                      eam partem Oceani quae est ad Hispaniam pertinet; spectat inter
occasum solis
                      et septentriones.]</seg>
<!-- ... -->
          </p>
<!-- ... -->
        </div>
<!-- Ch. 2 etc. ... -->
      </div>
    </body>
  </text>
  <standOff>
    <listPlace>
      <place xml:id="Gallia">
        <placeName xml:lang="la">Gallia</placeName>
        <placeName xml:lang="en">Gaul</placeName>
        <idno type="URI">https://pleiades.stoa.org/places/993</idno>
      </place>
<!-- ... -->
    </listPlace>
  </standOff>
  <TEI xmlns="http://www.tei-c.org/ns/1.0">
    <teiHeader>
<!-- Metadata applying just to the set of annotations -->
    </teiHeader>
    <standOff>
      <listAnnotation>
        <annotation xml:id="ann01"
          motivation="linking" target="#Gallia">
          <respStmt xml:id="ed">
```

```
          <resp>creator</resp>
          <persName>Fred Editor</persName>
        </respStmt>
        <revisionDesc>
          <change  status="created"
            when="2020-05-21T13:59:00Z"  who="#ed"/>
          <change  status="modified"
            when="2020-05-21T19:48:00Z"  who="#ed"/>
        </revisionDesc>
        <licence  target="http://creativecommons.org/licenses/by/3.0/"/>
<!-- Gallia in seg 1 -->
        <ptr  target="#string-range(c1p1s1,0,6)"/>
<!-- Galliae in seg 6 -->
        <ptr  target="#string-range(c1p1s6,19,7)"/>
       </annotation>
<!-- ... -->
     </listAnnotation>
   </standOff>
  </TEI>
</TEI>
```

## 17.12 Connecting Analytic and Textual Markup

In chapters *18. Simple Analytic Mechanisms* and *19. Feature Structures* and elsewhere, provision is made for analytic and interpretive markup to be represented outside of textual markup, either in the same document or in a different document. The elements in these separate domains can be connected, either with the pointing attributes *ana* (for *analysis*) and *inst* (for *instance*), or by means of <link> and <linkGrp> elements. Numerous examples are given in these chapters.

Another more specific form of annotation is available through the TEI <ruby> element and its children, described in detail in *3.4.2. Ruby Annotations*.

## 17.13 Module for Linking, Segmentation, and Alignment

The module described in this chapter makes available the following components:

**Module linking: Linking, segmentation, and alignment**

- *Elements defined*: ab alt altGrp anchor annotation join joinGrp link linkGrp listAnnotation seg standOff timeline when
- *Classes defined*: att.global.linking

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 18 Simple Analytic Mechanisms

This chapter describes a module for associating simple analyses and interpretations with text elements. We use the term *analysis* here to refer to any kind of semantic or syntactic interpretation which an encoder wishes to attach to all or part of a text. Examples discussed in this chapter include familiar linguistic categorizations (such as 'clause', 'morpheme', 'part-of-speech' etc.) and characterizations of narrative structure (such as 'theme', 'reconciliation' etc.). The mechanisms presented in this chapter are simpler but less powerful than those described in chapter *19. Feature Structures*.

Section *18.1. Linguistic Segment Categories* introduces elements which can be used to characterize text segments according to the familiar linguistic categories of *sentence* or *s-unit*, *clause*, *phrase*, *word*, *morpheme*, *character*, and *punctuation mark*. These elements represent special cases of the generic <seg> element described in section *17.3. Blocks, Segments, and Anchors*.

Section *18.2. Global Attributes for Simple Analyses* introduces an additional global attribute which allows passages of text to be associated with specialized elements representing their interpretation. These 'interpretative' elements ( and <interp>) are described in detail in section *18.3. Spans and Interpretations*. They allow the encoder to specify an analysis as a series of names and associated values,[109] each such pair being linked to one or more stretches of text, either directly, in the case of spans, or indirectly, in the case of interpretations.

Finally section *18.4. Linguistic Annotation* revisits the topic of linguistic analysis, and illustrates how these interpretative mechanisms may be used to associate simple linguistic analysis with text segments.

## 18.1 Linguistic Segment Categories

In this section we introduce specialized *linguistic segment category* elements which may be used to represent the segmentation of a text into the traditional linguistic categories of *sentence*, *clause*, *phrase*, *word*, *morpheme*, *characters*, and *punctuation marks*.

### 18.1.1 Words and Above

Although different languages have very different rules about what constitutes a 'word' or a 'sentence', these remain generally useful concepts. In this section we discuss elements provided for marking up linguistic units down to the word level, however defined.

`<s>` (s-unit) contains a sentence-like division of a text.

`<cl>` (clause) represents a grammatical clause.

`<phr>` (phrase) represents a grammatical phrase.

`<w>` (word) represents a grammatical (not necessarily orthographic) word.

As members of the `att.segLike` class, these elements all share the following attribute:

`att.segLike` provides attributes for elements used for arbitrary segmentation.

`@function` (function) characterizes the function of the segment.

They also share attributes from `att.typed`:

`att.typed` provides attributes that can be used to classify or subclassify elements in any way.

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.

`@subtype` (subtype) provides a sub-categorization of the element, if needed.

These elements are also all members of the `model.segLike` class, which is a subclass of `model.phrase`. They may thus appear anywhere that text is permitted within a document, when the module defined by this chapter is included in a schema.

The <w> and <pc> elements belong to the `att.linguistic` class, which supplies attributes that may be used for lightweight linguistic annotation (see section *18.4.2. Lightweight Linguistic Annotation* below):

`att.linguistic` provides a set of attributes concerning linguistic features of tokens, for usage within token-level elements, specifically <w> and <pc> in the analysis module.

`@lemma` provides a lemma (base form) for the word, typically uninflected and serving both as an identifier (e.g. in dictionary contexts, as a headword), and as a basis for potential inflections.

[109] Or, as they are widely known, *attribute-value pairs*; this term should not be confused, however, with XML attributes and their values, which are similar in concept but distinct in their formal definitions.

@*lemmaRef* provides a pointer to a definition of the lemma for the word, for example in an online lexicon.

@*pos* (part of speech) indicates the part of speech assigned to a token (i.e. information on whether it is a noun, adjective, or verb), usually according to some official reference vocabulary (e.g. for German: STTS, for English: CLAWS, for Polish: NKJP, etc.).

@*msd* (morphosyntactic description) supplies morphosyntactic information for a token, usually according to some official reference vocabulary (e.g. for German: STTS-large tagset; for a feature description system designed as (pragmatically) universal, see Universal Features).

@*join* when present, provides information on whether the token in question is adjacent to another, and if so, on which side.

Additionally, these elements also have access to the `att.lexicographic.normalized` class, which supplies the attributes *norm* and *orig*: the former for handling normalization/regularization at the word level, the latter providing the original form if the element content is modernized or regularized. Note that these attributes are a local (word-level) alternative to the robust mechanism that uses the <choice>, <orig>, and <reg> elements, discussed in section *3.5.2. Regularization and Normalization* and in chapter *13. Critical Apparatus*. The <choice>-based mechanism is the default descriptive device, while the *norm* and *orig* attributes are used to handle a subset of normalizations in linguistic contexts where a single sequence of tokens is a priority, for example in historical corpora subject to linguistic analysis. It needs to be stressed that the simplified attribute-based mechanism is not meant to be used for editorial interventions. [110]

The <s> element may be used simply to segment a text end-to-end into a series of non-overlapping segments, referred to here and elsewhere as *s-units*, or *sentences*.

```
<p>
  <s>Nineteen fifty-four, when I was eighteen years old,
     is held to be a crucial turning point in the history of
     the Afro-American — for the U.S.A. as a whole — the
     year segregation was outlawed by the U.S. Supreme Court.</s>
  <s>It was also a crucial year for me because on June 18,
     1954, I began serving a sentence in state prison for
     possession of marijuana.</s>
</p>
```

The <s> element is more restricted both in its content and its usage than the generic <seg> element. The <seg> unit may contain anything which can appear within a paragraph: thus it may be used to enclose members of the `model.inter` class (such as <bibl> or <list>) as well as other phrase elements; the <s> unit may only contain phrase-level elements or text. Also, unlike <seg> elements, <s> elements should not be nested within each other.[111] The <seg> element is intended for use as a generic segmentation element, the specific function of which may be indicated by its *type* attribute; the other members of the class are more specialized. Thus, the <s>, <cl>, and <phr> elements may be thought of as equivalent to <seg type="s-unit">, <seg type="clause"> and <seg type="phrase">, respectively, but with the above-mentioned restrictions.

The <s> element may be further subdivided into *clauses*, marked with the <cl> element, as in the following example:

```
<p>
  <s>
    <cl>It was about the beginning of September, 1664,
    <cl>that I, among the rest of my neighbours,
           heard in ordinary discourse
      <cl>that the plague was returned again to Holland; </cl>
      </cl>
```

[110] The `att.lexicographic.normalized` class is also used in dictionary entries, as discussed in chapter *10. Dictionaries*.

[111] Neither this constraint, nor the requirement that the whole of the text be segmented by <s> elements is required by the TEI Guidelines.

```
    </cl>
    <cl>for it had been very violent there, and particularly at
        Amsterdam and Rotterdam, in the year 1663, </cl>
    <cl>whither, <cl>they say,</cl> it was brought,
    <cl>some said</cl> from Italy, others from the Levant, among some goods
    <cl>which were brought home by their Turkey fleet;</cl>
    </cl>
    <cl>others said it was brought from Candia;
        others from Cyprus. </cl>
  </s>
  <s>
    <cl>It mattered not <cl>from whence it came;</cl>
    </cl>
    <cl>but all agreed <cl>it was come into Holland again.</cl>
    </cl>
  </s>
</p>
```

Clauses may be further divided into <phr> elements in the same way. A text may be segmented directly into clauses, or into phrases, with no need to include segmentation at a higher level as well.

For verse texts, the overlapping of metrical and syntactic structure requires that special care be given to representing both using an element hierarchy. One simple approach is to split the syntactic phrases into fragments when they cross verse boundaries, reuniting them with the *part* attribute:

```
<div  type="stanza">
  <l>
    <cl  part="I">Tweedledum and Tweedledee</cl>
  </l>
  <l>
    <cl  part="F">Agreed to have a battle;</cl>
  </l>
  <l>
    <cl  part="I">For Tweedledum said <cl  part="I">Tweedledee</cl>
    </cl>
  </l>
  <l>
    <cl  part="F">
      <cl  part="F">Had spoiled his nice new rattle.</cl>
    </cl>
  </l>
</div>
<div  type="stanza">
  <l>
    <cl  part="I">Just then flew down a monstrous crow,</cl>
  </l>
  <l>
    <cl  part="F">As black as a tar barrel;</cl>
  </l>
  <l>
    <cl  part="I">Which frightened both the heroes so,</cl>
  </l>
  <l>
    <cl  part="F">
      <cl>They quite forgot their quarrel.</cl>
    </cl>
```

```
  </l>
</div>
```

Another approach is to use the *next* and *prev* attributes defined in the additional module for linking (chapter *17. Linking, Segmentation, and Alignment*):

```
<l>
  <cl next="#c5" xml:id="c3" part="I">For Tweedledum said
  <cl next="#c6" xml:id="c4" part="I">Tweedledee</cl>
  </cl>
</l>
<l>
  <cl prev="#c3" xml:id="c5" part="F">
    <cl prev="#c4" xml:id="c6" part="F">Had spoiled his nice new rattle.</cl>
  </cl>
</l>
```

Other methods are also possible; for discussion, see chapter *21. Non-hierarchical Structures*.

The *type* attribute on linguistic segment categories can be used to provide additional interpretative information about the category. The *function* attribute on the <cl> and <phr> elements can be used to provide additional information about the function of the category. Legal values for these two attributes are not defined by these Guidelines, but should be documented in the <segmentation> element of the <encodingDesc> element within the document's header. A general approach to the encoding of linguistic categories for parts of a text is discussed in section *18.4. Linguistic Annotation* below.

Using traditional terminology, these attributes provide a convenient way of specifying, for example, that the clause *from whence it came* is a relative clause modifying another, or that the phrase *by the U.S. Supreme Court* is a prepositional post-modifier:

```
<cl>It mattered not
<cl type="relative"
   function="clause_modifier">from whence it came;</cl>
</cl>
```

```
<phr type="NP">the year segregation</phr>
<phr>was outlawed</phr>
<phr type="PP"
 function="postmodifier-agent">by the U.S. Supreme Court.</phr>
```

Segmentation into clauses and phrases can, of course, be combined. Such detailed encodings as the following may require careful formatting if they are to be easily readable however.

```
<p>
  <s>
    <cl type="finite-declarative"
     function="independent">
     <phr type="NP" function="subject">Nineteen fifty-four,
     <cl type="finite-relative-declarative"
        function="appositive">when <phr type="NP" function="subject">I</phr>
        <phr type="VP" function="predicate">was eighteen years old</phr>
      </cl>
```

```
    </phr>,
  <phr  type="VP"  function="predicate">
     <phr  type="V"  function="verb-main">is held</phr>
     <phr  type="NP"  function="complement">
      <cl  type="nonfinite"
       function="predicate-nom.">
       <phr  type="V"  function="copula">to be</phr>
       <phr  type="NP"
        function="predicate-nom.">a crucial turning point
       <phr  type="PP"
          function="postmodifier">in
        <phr  type="NP"  function="prep.obj.">the history
          <phr  type="PP"
             function="postmodifier">of the Afro-American</phr>
          </phr>
        </phr>
                 –
       <phr  type="PP"
          function="postmodifier-appositive">for
        <phr  type="NP"  function="prep.obj.">the U.S.A.
          <phr  type="PP"
             function="postmodifier">as a whole</phr>
          </phr>
        </phr>
       </phr>
              –
      <phr  type="NP"
         function="appositive-predicate-nom.">the year
        <cl  type="finite-relative"
           function="adjectival">
           <phr  type="NP"  function="subject">segregation</phr>
           <phr  type="VP"  function="predicate">
            <phr  type="V"  function="verb-main">was outlawed</phr>
            <phr  type="PP"
              function="postmodifier">by the U.S. Supreme Court</phr>
           </phr>
          </cl>
        </phr>
       </cl>
     </phr>
   </phr>.</cl>
</s>
<s>
  <cl  type="finite-declarative"
    function="independent">
    <phr  type="NP"  function="subject">It</phr>
    <phr  type="VP"  function="predicate">
      <phr  type="V"  function="verb-main">was</phr>
          also
    <phr  type="NP"
       function="predicate-nom.">a crucial year for me</phr>
    </phr>
    <cl  type="declarative-finite"
      function="dependent-causative">because
    <phr  type="PP"
       function="sentence_adverb">on June 18, 1954</phr>,
    <phr  type="NP"  function="subject">I</phr>
```

```
      <phr type="VP" function="predicate">
        <phr type="V" function="verb-main">began serving</phr>
        <phr type="NP" function="complement">a sentence in state prison
        <phr type="PP" function="complement">for possession of marijuana</phr>
        </phr>
      </phr>
     </cl>
    </cl>
  </s>.
</p>
```

This style of markup may introduce spurious new lines and blanks into the text. If the original layout is important, it should be explicitly encoded, using such facilities as the <lb> element, the global *rend* or *rendition* attributes, etc.

The <w>, <m>, and <c> elements are identical in meaning to the <seg> element with a *type* attribute of 'w', 'm', or 'c' respectively, and may occur wherever <seg> is permitted to occur. However, their content is more constrained than <seg>: for example, the <w> element should only contain <w>, <m>, <c> elements or <pc> elements, or plain text; the <m> element should contain only <c> or <pc> elements or plain text; both the <c> and <pc> elements should contain only plain text, most often only a single character or a sequence of graphemes to be treated as a single character. Consequently, while these more specific elements can be translated directly into typed <seg> elements, the reverse is not necessarily the case.

The restriction on the content of the <w> element in particular requires that a certain care must be exercised when using it, especially in relation to the use of other tags that one may think of as *word level*, but which are in fact defined as *phrase level*. Consider the problem of segmenting an occurrence of the <mentioned> element as a word.

```
<mentioned>grandiloquent</mentioned>
```

The first of the following two encodings is legitimate; the second is not, since the <mentioned> element is not part of the content model of the <w> element:

```
<mentioned>
  <w>grandiloquent</w>
</mentioned>
```

```
<w>
  <mentioned>grandiloquent</mentioned>
</w>
```

On the other hand, both of the following encodings *are* legitimate:

```
<mentioned>
  <phr>grandiloquent speech</phr>
</mentioned>
```

```
<phr>
  <mentioned>grandiloquent speech</mentioned>
</phr>
```

The first encoding describes the citing of a phrase. The second describes a phrase which consists of something mentioned.

The <w> element carries additional attributes which may be of use in many indexing or analytic applications. The *lemma* attribute may be used to specify the *lemma*, that is the head- or uninflected form of an inflected verb or noun, for example:

```
<s xml:lang="la">
  <w lemma="timeo">timeo</w>
  <w lemma="danaii">Danaos</w>
  <w lemma="et">et</w>
  <w lemma="donum">dona</w>
  <w lemma="fero">ferentes</w>
</s>
```

In some situations it may be more convenient to use the *lemmaRef* pointer attribute than to supply an explicit uninflected form. This attribute assumes the existence of a list of uninflected forms, for example in an online lexicon, with which individual w entries can be associated using the usual TEI pointer mechanisms. Assuming that a standardized lexicon for Latin is available at the location `http://lexicon.org/latin.xml`, we might for example revise the above example as:

```
<s xml:lang="la">
  <w lemmaRef="http://lexicon.org/latin.xml#timeo">timeo</w>
  <w lemmaRef="http://lexicon.org/latin.xml#danaii">Danaos</w>
<!-- ... -->
</s>
```

### 18.1.2 Below the Word Level

It is sometimes helpful to markup explicitly sub-word components such as morphemes, characters, or punctuation.

`<m>` (morpheme) represents a grammatical morpheme.

`<c>` (character) represents a character.

`<pc>` (punctuation character) contains a character or string of characters regarded as constituting a single punctuation mark.

The <m> element is used to mark up morphologically identified segmentation below the word level. Analogous to the *lemma* attribute for <w>, there is a *baseForm* attribute for the <m> element, which may be used to indicate the 'base form' of an inflected morpheme; where appropriate, <m> elements may also be organized hierarchically:

```
<w type="adjective">
  <m type="base">
    <m type="prefix" baseForm="con">com</m>
    <m type="root">fort</m>
  </m>
  <m type="suffix">able</m>
</w>
```

The distinction between <m> and <w> is provided as a convenience only; it may not be appropriate for all linguistic theories, nor is it meaningful in all languages. The intention is to provide a means for those cases where it is considered helpful to distinguish lexical from sub-lexical tokens, to complement the more general mechanism already provided by the <seg> element, using which the above example could alternatively be marked up as follows:

```
<seg type="adjective">
  <seg type="base">
    <seg type="prefix">com</seg>
    <seg type="morph">fort</seg>
  </seg>
  <seg type="suffix">able</seg>
</seg>
```

See section *18.4.2. Lightweight Linguistic Annotation* for an alternative to using *type* in such contexts.

There is a substantial linguistic difference between characters like letters or diacritics and punctuation marks. The former are used to construct meaningful units like morphemes or words. The latter are functionally independent units acting at the level of syntactic units. A word may consist of a single letter (for example 'I' in English), but this does not mean that we should use <c> instead of <w> to mark it up.

The <c> (character) element should be used to mark up any non-lexical character, whether this appears within a word, or outside it. In the following example, the encoder wishes to indicate that the letters are not to be regarded as words:

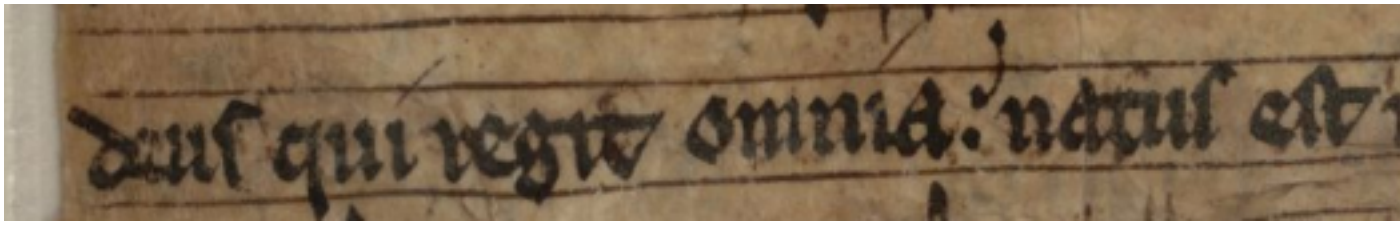
```
<phr>
  <c>M</c>
  <c>O</c>
  <c>A</c>
  <c>I</c>
  <w>doth</w>
  <w>sway</w>
  <w>my</w>
  <w>life</w>
</phr>
```

The <c> element may be used for individual characters occurring within a <w> or <m> element which it is desired to distinguish for some reason, as in the following examples:

```
<m baseForm="not">
  <c>n</c>
  <c type="punct">'</c>
  <c>t</c>
</m>
```

This encoding represents the constituents of a common abbreviation, but does not indicate that it is in fact an abbreviation; the <am> element (*12.3.1.2. Abbreviation and Expansion*) may be preferred for the latter purpose. Generally speaking, the use of <c> use to mark non-lexical punctuation marks is deprecated, since the <pc> element is provided specifically to distinguish these.

The <pc> (punctuation character) element should be used to mark up characters which are specifically regarded as providing punctuation, rather than constituting parts of a word. It may be particularly useful when transcribing older written materials, in which an encoding of the original punctuation may be useful for interpretive or analytic purposes, in much the same way as an encoding of the original orthography may be. For example, in the following extract from a Bodleian Library musical manuscript two different punctuation marks are used to distinguish kinds of pause in the text. The *punctus elevatus* (which resembles an inverted semicolon) is not a Unicode character, but may still be encoded using the <g> element. As further described in chapter *5. Characters, Glyphs, and Writing Modes*, this element points to a definition for the intended character which may be stored either locally or elsewhere.



```
deus qui regis omnia
<pc>
  <g  ref="#pelev">;</g>
</pc> natus est in bethlehem
<pc>.</pc>o <pc>.</pc> mira gratia...

<!-- elsewhere -->
<char  xml:id="pelev">
<!-- definition of the punctus elevatus character -->
</char>
```

The <pc> element carries special attributes to record analyses of the functional behaviour or classification of the punctuation mark it contains. The *unit* attribute may be used, as on the <milestone> element to name the kind of unit which the punctuation mark delimits, for example a paragraph or section. The *pre* attribute may be used to indicate whether the punctuation precedes or follows the unit it delimits. The *force* attribute indicates the strength of the association between the punctuation mark and its adjacent word.

In the following example, the paragraph marker (¶) has been tagged as a strong punctuation mark, preceding the unit it marks, which is named 'para':

```
<p>
  <pc  unit="para"  force="strong"  pre="true">¶</pc>Incipit...
</p>
```

A similar encoding can be used for hyphenation:

```
A fire<pc  force="strong">-</pc>proof vest is recom<pc  force="weak">-</pc><lb/>
mended.
```

Refer to *3.2.2. Hyphenation* for a discussion of the motivations for explicitely recording the presence of hyphens.

The <w>, <m>, <c>, and <pc> elements can be used together to give a fairly detailed low-level grammatical analysis of text. For example, consider the following segmentation of the English S-unit *I didn't do it.*

```
<w>I</w>
<w>
  <m  baseForm="do">did</m>
  <m>n't</m>
</w>
<w  lemma="do">do</w>
<w>it</w>
<pc>.</pc>
```

This segmentation, crude as it is, succeeds in representing the idea that *did* occurring as a morphological component of the word *didn't* has something in common with the word <do>. A further advantage of segmenting

the text down to this level is that it becomes relatively simple to associate each such segment with a more detailed formal analysis, for example by providing a baseform, or morphological analysis at whichever level is appropriate. This matter is taken up in detail in section *18.4. Linguistic Annotation*.

## 18.2 Global Attributes for Simple Analyses

When the module described by this chapter is selected, an additional attribute is defined for all elements:

`att.global.analytic` provides additional global attributes for associating specific analyses or interpretations with appropriate portions of a text.

- `@ana` (analysis) indicates one or more elements containing interpretations of the element on which the *ana* attribute appears.

The *ana* attribute may be specified for any element. Its effect is to associate the element with one or more others representing an analysis or interpretation of it. Its target should be one of the elements described in the section *18.3. Spans and Interpretations* below, or some other interpretative element such as <note>, on which see section *3.9. Notes, Annotation, and Indexing* or <fs>, on which see chapter *19. Feature Structures*. If a hierarchical form of classification is desired then it may point to <category> element at a suitable level in a <taxonomy> see *2.3.7. The Classification Declaration*.

## 18.3 Spans and Interpretations

The simplest mechanisms for attaching analytic notes in some structured vocabulary to particular passages of text are provided by the  and <interp> elements, and their associated grouping elements <spanGrp> and <interpGrp>.

`<span>` associates an interpretative annotation directly with a span of text.

`<spanGrp>` (span group) collects together span tags.

`<interp>` (interpretation) summarizes a specific interpretative annotation which can be linked to a span of text.

`<interpGrp>` (interpretation group) collects together a set of related interpretations which share responsibility or type.

These elements are all members of the class `att.interpLike`, and thus share the following attribute:

`att.interpLike` provides attributes for elements which represent a formal analysis or interpretation.

- `@inst` (instances) points to instances of the analysis or interpretation represented by the current element.

They also inherit the following attributes from `att.global.responsibility`:

`att.global.responsibility` provides attributes indicating the agent responsible for some aspect of the text, the markup or something asserted by the markup, and the degree of certainty associated with it.

- `@cert` (certainty) signifies the degree of certainty associated with the intervention or interpretation.
- `@resp` (responsible party) indicates the agency responsible for the intervention or interpretation, for example an editor or transcriber.

The *type* attribute of the  and <interp> elements may be used to indicate that the annotations are of specific types, for example thematic or structural. The annotation itself is supplied as the content of the  or <interp> element. In the case of the  element, the span of text being annotated is indicated by values of the *from*, *to* or *target* attributes, used in combination as follows. If only the *from* attribute is supplied, then the span is coterminous with the element indicated by its value; if both *from* and *to* are supplied, the span runs from the start of the element indicated by the *from* attribute up to the end of the element indicated by the *to* attribute; if the *target* attribute is used, the span is defined by aggregating the contents of the (possibly non-contiguous) elements pointed to by its values. It is an error to supply only the *to* attribute; to supply more than one pointer value for either *to* or *from* attributes; or to supply either of these in conjunction with the *target* attribute. In the case of <interp> (see below), the span is indicated by a pointer from a <link> element or some similar mechanism. The *resp* attribute indicates the annotator responsible for this annotation.

The  element provides a simple way of indicating such features as phrasal verbs in a linguistic analysis, as in this example:

```
<s>
  <w>What</w>
  <w>did</w>
  <w>you</w>
  <w  xml:id="mk01">make</w>
  <w  xml:id="up01">up</w>
</s>
<span  from="#mk01"  to="#up01">phrasal verb "make up"</span>
```

Here the two components of the span follow each other, so the *to* and *from* attributes may be used. The same effect could however be achieved by using the *target* attribute:

```
<s>
  <w>What</w>
  <w>did</w>
  <w>you</w>
  <w  xml:id="mk02">make</w>
  <w  xml:id="up02">up</w>
</s>
<span  target="#mk02 #up02">phrasal verb "make up"</span>
```

This second approach might be cumbersome if the number of components to be combined is very large. It is however essential if the components do not follow each other, as in this example:

```
<s>
  <w>Did</w>
  <w>you</w>
  <w  xml:id="mk03">make</w>
  <w>it</w>
  <w  xml:id="up03">up</w>
</s>
<span  target="#mk03 #up03">phrasal verb "make up"</span>
```

The  element can be used for any kind of annotation. In this example it is used in a narratological analysis:

```
<p  xml:id="MaQp1s2p114">
  <s  xml:id="MaQp1s2p114s1">There was certainly a definite point at which the
     thing began.</s>
  <s  xml:id="MaQp1s2p114s2">It was not; then it was suddenly inescapable,
     and nothing could have frightened it away.</s>
  <s  xml:id="MaQp1s2p114s3">There was a slow integration, during which she,
     and the little animals, and the moving grasses, and the sun-warmed
     trees, and the slopes of shivering silvery mealies, and the great
     dome of blue light overhead, and the stones of earth under her feet,
     became one, shuddering together in a dissolution of dancing
     atoms.</s>
  <s  xml:id="MaQp1s2p114s4">She felt the rivers under the ground forcing
     themselves painfully along her veins, swelling them out in an
     unbearable pressure; her flesh was the earth, and suffered growth
     like a ferment; and her eyes stared, fixed like the eye of the
     sun.</s>
  <s  xml:id="MaQp1s2p114s5">Not for one second longer (if the terms for time
     apply) could she have borne it; but then, with a sudden movement
```

```
    forwards and out, the whole process stopped; and <emph  rend="italic">that</emph>
was <soCalled  rend="dquo">the
       moment</soCalled> which it was impossible to remember
    afterwards.</s>
  <span  from="#MaQp1s2p114s3"
    to="#MaQp1s2p114s5">the moment</span>
  <s  xml:id="MaQp1s2p114s6">For during that space of time (which was
    timeless) she understood quite finally her smallness, the
    unimportance of humanity.</s>
</p>
```

The  element may, as in this example, be placed in the text near the textual span it is associated with. Alternatively, it may be placed elsewhere in the same or a different document. Where several  or <interp> elements share the same attributes, for example having the same responsibility or type, it may be convenient to group them within a <spanGrp> or <interpGrp> element as follows:

```
<spanGrp  resp="#DTL">
  <span  from="#MaQp1s2p114s3"
    to="#MaQp1s2p114s5">the moment</span>
<!-- other spans identified by DTL here -->
</spanGrp>
```

Spans may also be used to represent structural divisions within a narrative, particularly when these do not coincide with the structure implied by the element structure. Consider the following narrative:

> Sigmund, the son of Volsung, was a king in Frankish country. Sinfiotli was the eldest of his sons, the second was Helgi, the third Hamund. Borghild, Sigmund's wife, had a brother named — But Sinfiotli, her stepson, and — both wooed the same woman and Sinfiotli killed him over it.[112] And when he came home, Borghild asked him to go away, but Sigmund offered her weregild, and she was obliged to accept it. At the funeral feast Borghild was serving beer. She took poison, a big drinking horn full, and brought it to Sinfiotli. When Sinfiotli looked into the horn, he saw that poison was in it, and said to Sigmund 'This drink is cloudy, old man.' Sigmund took the horn and drank it off. It is said that Sigmund was hardy and that poison did him no harm, inside or out. And all his sons could tolerate poison on their skin. Borghild brought another horn to Sinfiotli, and asked him to drink, and everything happened as before. And a third time she brought him a horn, and reproachful words as well, if he didn't drink from it. He spoke again to Sigmund as before. He said 'Filter it through your mustache, son!' Sinfiotli drank it off and at once fell dead.
>
> Sigmund carried him a long way in his arms and came to a long, narrow fjord, and there was a small boat there and a man in it. He offered to ferry Sigmund over the fjord. But when Sigmund carried the body out to the boat, it was fully laden. The man said Sigmund should go around the fjord inland. The man pushed the boat out and then suddenly vanished.
>
> King Sigmund lived a long time in Denmark in the kingdom of Borghild, after he married her. Then he went south to Frankish lands, to the kingdom he had there. Then he married Hiordis, the daughter of King Eylimi. Their son was Sigurd. King Sigmund fell in a battle with the sons of Hunding. And then Hiordis married Alf, the son of King Hialprec. Sigurd grew up there as a boy.
>
> Sigmund and all his sons were tall and outstanding in their strength, their growth, their intelligence, and their accomplishments. But Sigurd was the most outstanding of all, and everyone who knows about the old days says he was the most outstanding of men and the noblest of all the warrior kings.

A structural analysis of this text, dividing it into narrative units in a pattern shared with other texts from the same literature, might look like this:

[112] The rule marks spaces left for the missing name in the manuscript.

```
<p xml:id="P1">
  <s xml:id="S1">Sigmund ... was a king in Frankish country.</s>
  <s xml:id="S2">Sinfiotli was the eldest of his sons.</s>
  <s xml:id="S3">Borghild, Sigmund's wife, had a brother ...</s>
  <s xml:id="S4A">But Sinfiotli ... wooed the same woman</s>
  <s xml:id="S4B">and Sinfiotli killed him over it.</s>
  <s xml:id="S5">And when he came home, ... she was obliged to accept it.</s>
  <s xml:id="S6">At the funeral feast Borghild was serving beer.</s>
  <s xml:id="S7">She took poison ... and brought it to Sinfiotli.</s>
  <s xml:id="S17">Sinfiotli drank it off and at once fell dead.</s>
  <anchor xml:id="EOS17"/>
</p>
<p xml:id="P2">Sigmund carried him a long way in his arms ... </p>
<p xml:id="P3">King Sigmund lived a long time in Denmark ... </p>
<p xml:id="P4">Sigmund and all his sons were tall ... </p>
<spanGrp resp="#TMA"
  type="narrative-structure">
  <span from="#S1" to="#S3">introduction</span>
  <span from="#S4A">conflict</span>
  <span from="#S4B">climax</span>
  <span from="#S5" to="#S17">revenge</span>
  <span from="#EOS17">reconciliation</span>
  <span from="#P2" to="#P4">aftermath</span>
</spanGrp>
```

Note the use of an empty <anchor> element to provide a target for the 'reconciliation' unit which is normally part of the narrative pattern but which is not realized in the text shown.

The same analysis may be expressed with the <interp> element instead of the  element; this element provides attributes for recording an interpretive category and its value, as well as the identity of the interpreter, but does not itself indicate which passage of text is being interpreted; the same interpretive structures can thus be associated with many passages of the text. The association between text passages and <interp> elements should be made either by pointing from the text to the <interp> element with the *ana* attribute defined in section *18.2. Global Attributes for Simple Analyses*, or by pointing at both text and interpretation from a <link> element, as described in chapter *17. Linking, Segmentation, and Alignment*.

To encode the first example above using <interp>, it is necessary to create a text element which contains—or corresponds to—the third, fourth, and fifth orthographic sentences (S-units) in the paragraph. This can be done either with the <seg> element, described in *17.3. Blocks, Segments, and Anchors*, or the <join> element, described in *17.7. Aggregation*. The resulting element can then be associated with the <interp> element using the *ana* attribute described in section *18.2. Global Attributes for Simple Analyses*. We illustrate using the <seg> element.

```
<p xml:id="MarQp1s2p114">
  <s xml:id="MarQp1s2p114s1">There was certainly a definite point ... </s>
  <s xml:id="MarQp1s2p114s2">It was not; then it was suddenly inescapable ... </s>
  <seg xml:id="MarQp1s2p114s3-5"
    ana="#moment">
    <s xml:id="MarQp1s2p114s3">There was a slow integration ... </s>
    <s xml:id="MarQp1s2p114s4">She felt the rivers under the ground ... </s>
    <s xml:id="MarQp1s2p114s5">Not for one second longer ... </s>
  </seg>
  <s xml:id="MarQp1s2p114s6">For during that space of time ... </s>
</p>
<interp xml:id="moment">the moment</interp>
```

The second example above can be recoded using <interp> and <interpGrp> tags in a similar manner. The

interpretation itself can be expressed in an <interpGrp> element, which would replace the <spanGrp> in the example shown above:

```
<interpGrp  resp="#TMA"
  type="structuralunit">
  <interp  xml:id="INTRO">introduction</interp>
  <interp  xml:id="CONFLICT">conflict</interp>
  <interp  xml:id="CLIMAX">climax</interp>
  <interp  xml:id="REVENGE">revenge</interp>
  <interp  xml:id="RECONCIL">reconciliation</interp>
  <interp  xml:id="AFTERM">aftermath</interp>
</interpGrp>
```

Any of these <interp> elements may be linked to the text either by means of the *ana* attribute, or by means of <link> elements. Using the *ana* attribute (on <seg> elements introduced specifically for this purpose), the text would be encoded as follows:

```
<p  xml:id="PP1">
  <seg  xml:id="SS1-SS3"  ana="#INTRO">
    <s  xml:id="SS1">Sigmund ... was a king in Frankish country.</s>
    <s  xml:id="SS2">Sinfiotli was the eldest of his sons.</s>
    <s  xml:id="SS3">Borghild, Sigmund's wife, had a brother ... </s>
  </seg>
  <s  xml:id="SS4A"  ana="#CONFLICT">But Sinfiotli ... wooed the same woman</s>
  <s  xml:id="SS4B"  ana="#CLIMAX">and Sinfiotli killed him over it.</s>
  <seg  xml:id="SS5-SS17"  ana="#REVENGE">
    <s  xml:id="SS5">And when he came home, ... she was obliged to accept it.</s>
    <s  xml:id="SS6">At the funeral feast Borghild was serving beer.</s>
    <s  xml:id="SS17">Sinfiotli drank it off and at once fell dead.</s>
  </seg>
</p>
<anchor  xml:id="NIL1"  ana="#RECONCIL"/>
<p  xml:id="PP2">Sigmund carried him a long way in his arms ... </p>
<p  xml:id="PP3">King Sigmund lived a long time in Denmark ... </p>
<p  xml:id="PP4">Sigmund and all his sons were tall ... </p>
<join  xml:id="PP2-PP4"
  target="#PP2 #PP3 #PP4"  ana="#AFTERM"/>
```

The linkage may also be accomplished using a <linkGrp> element, whose content is a set of <link> elements which point to each interpretive element and its corresponding text unit. This method does not require the use of the *ana* attribute on the text units.

```
<linkGrp  targFunc="interpretation text">
  <link  target="#INTRO #SS1-SS3"/>
  <link  target="#CONFLICT #SS4A"/>
  <link  target="#CLIMAX #SS4B"/>
  <link  target="#REVENGE #SS5-SS17"/>
  <link  target="#RECONCIL #NIL1"/>
  <link  target="#AFTERM #PP2-PP4"/>
</linkGrp>
```

One obvious advantage of using <interp> rather than  elements for the Sigmund text is that the <interp> elements can be reused for marking up other texts in the same document, whereas the  elements cannot. On the other hand, the use of <interp> elements may require the creation of special text elements not otherwise needed (e.g. the <seg> and the <join> in the revised encoding of the text), whereas the use of  elements does not.

## 18.4 Linguistic Annotation

By *linguistic annotation* we mean here any annotation determined by an analysis of linguistic features of the text, excluding as borderline cases both the formal structural properties of the text (e.g. its division into chapters or paragraphs) and descriptive information about its context (the circumstances of its production, its genre or medium). The structural properties of any TEI-conformant text should be represented using the structural elements discussed elsewhere in this chapter and in chapters *3. Elements Available in All TEI Documents*, *4. Default Text Structure*, *6. Verse*, *7. Performance Texts*, *8. Transcriptions of Speech*, *10. Dictionaries*, and *16. Language Corpora.* The contextual properties of a TEI text are fully documented in the TEI header, which is discussed in chapter *2. The TEI Header*, and in section *16.2. Contextual Information.*

Other forms of linguistic annotation may be applied at a number of levels in a text. A code (such as a word-class or part-of-speech code) may be associated with each word or token, or with groups of such tokens, which may be continuous, discontinuous, or nested. A code may also be associated with relationships (such as cohesion) perceived as existing between distinct parts of a text. The codes themselves may stand for discrete and non-decomposable categories, or they may represent highly articulated bundles of textual features. Their function may be to place the annotated part of the text somewhere within a narrowly linguistic or discoursal domain of analysis, or within a more general semantic field, or any combination drawn from these and other domains.

The manner by which such annotations are generated and attached to the text may be entirely automatic, entirely manual or a mixture. The ease and accuracy with which analysis may be automated may vary with the level at which the annotation is attached. The method employed should be documented in the <interpretation> element within the encoding description of the TEI header, as described in section *2.3.3. The Editorial Practices Declaration.* Where different parts of a language corpus have used different annotation methods, the *decls* attribute may be used to indicate the fact, as further discussed in section *16.3. Associating Contextual Information with a Text.*

### 18.4.1 Linguistic Annotation by Means of Generic TEI Devices

As one example of such types of analysis, consider the following sentence, taken from the Lancaster/IBM Treebank Project (**?**).

> The victim's friends told police that Kruger drove into the quarry and never surfaced.

Our discussion focuses on the way that this sentence might be analysed using the CLAWS system developed at the University of Lancaster but exactly the same principles may be applied to a wide variety of other systems.[113] Output from the system consists of a segmented and tokenized version of the text, in which word class codes have been associated with each token. CLAWS offers outputs in a variety of non-XML and XML formats: for example, the simplest format for the sample sentence would be:

```
  The_ AT0 victim_ NN1 's_ POS friends_ NN2 told_ VVD police_ NN2 that_ CJT Kruger_
NP0
drove_ VVD into_ PRP the_ AT0 quarry_ NN1 and_ CJC never_ AV0 surfaced_ VVD
```

This may be easily transformed into an equivalent TEI XML representation:

```
<s>
  <w  ana="#AT0">The </w>
  <w  ana="#NN1">victim</w>
  <w  ana="#POS">'s</w>
  <w  ana="#NN2">friends </w>
  <w  ana="#VVD">told </w>
  <w  ana="#NN2">police </w>
  <w  ana="#CJT">that </w>
  <w  ana="#NP0">Kruger </w>
```

[113] For the word-class tagging method used by CLAWS see **?**; For an overview of the system see **?**. The example sentence was processed using an online version of the CLAWS tagger at `http://ucrel.lancs.ac.uk/claws/`

```
  <w  ana="#VVD">drove </w>
  <w  ana="#PRP">into </w>
  <w  ana="#AT0">the </w>
  <w  ana="#NN1">quarry </w>
  <w  ana="#CJC">and </w>
  <w  ana="#AV0">never </w>
  <w  ana="#VVD">surfaced</w>
</s>
```

Although the names used for the attribute values here may have some significance for the human reader (AT0 for *article*, NN1 for *singular noun*, NN2 for *plural noun*, etc.) they are arbitrary codes, used in this case as pointers to other elements which define their significance more precisely. If the codes are considered to be *atomic*, then the <interp> element described in section *18.3. Spans and Interpretations* might be used to supply brief definitions in the header:

```
<interpGrp  type="POS">
  <interp  xml:id="AT0">Definite article</interp>
  <interp  xml:id="AV0">Adverb</interp>
  <interp  xml:id="CJC">Conjunction</interp>
  <interp  xml:id="CJT">Relative that</interp>
  <interp  xml:id="NN1">Noun singular</interp>
  <interp  xml:id="NN2">Noun plural</interp>
  <interp  xml:id="NP0">Proper noun</interp>
  <interp  xml:id="POS">Genitive marker</interp>
  <interp  xml:id="PRP">Preposition</interp>
  <interp  xml:id="VVD">Verb past tense</interp>
</interpGrp>
```

If the codes are considered to be compositional (for example that NN1 and NN2 have something in common, namely their *noun-ness*, which they do not share with, say, VVD), then this compositionality may be most clearly expressed using a mechanism based on the <fs> element defined in chapter *19. Feature Structures*.

This approach requires the text to be fully segmented, using the linguistic segment elements described in section *18.1. Linguistic Segment Categories*, so that the scope of the *ana* attribute used to point to each interpretation is clearly defined. A further analysis into phrase and clause elements can be superimposed on the word and morpheme tagging in the preceding illustration. For example, CLAWS provides the following constituent analysis of the sample sentence (the word class codes have been deleted):

```
  [N [G The victim's G] friends N] [V told [N police N] [Fn that
[N Krueger N] [V [V& drove [P into [N the quarry N]P]V&] and
[V+ never surfaced V+]V]Fn]V]
```

Treating the labels on the brackets as phrase or clause interpretations, this analysis of the structure of the example sentence can be combined with the word class analysis and represented as follows (the symbol V&”/> representing the first part of a coordinate phrase, has been replaced by V1, and V+, representing the second part, has been replaced by V2).

```
<s  type="sentence">
  <phr  ana="#n">
    <phr  ana="#gn">
      <w  ana="#AT0">The</w>
```

```
      <w  ana="#NN1">victim</w>
      <m  ana="#POS">'s</m>
    </phr>
    <w  ana="#NN2">friends</w>
  </phr>
  <phr  ana="#v">
    <w  ana="#VVD">told</w>
    <phr  ana="#n">
      <w  ana="#NN2">police</w>
    </phr>
    <cl  ana="#fn">
      <w  ana="#CJT">that</w>
      <phr  ana="#n">
        <w  ana="#NP0">Krueger</w>
      </phr>
      <phr  ana="#v">
        <phr  ana="#v1">
          <w  ana="#VVD">drove</w>
          <phr  ana="#pr">
            <w  ana="#PRP">into</w>
            <phr  ana="#n">
              <w  ana="#AT0">the</w>
              <w  ana="#NN1">quarry</w>
            </phr>
          </phr>
        </phr>
        <w  ana="#CJC">and</w>
        <phr  ana="#v2">
          <w  ana="#AV0">never</w>
          <w  ana="#VVD">surfaced</w>
        </phr>
      </phr>
    </cl>
  </phr>
  <c  ana="#pun">.</c>
</s>
```

This approach requires the definition of further <interp> (or <fs>) elements to provide targets for the pointers used to represent the constituent analysis:

```
<interpGrp  type="constituentFunction">
  <interp  xml:id="v2">coordinate continuation</interp>
  <interp  xml:id="v">verbal</interp>
  <interp  xml:id="no">nominal</interp>
  <interp  xml:id="gn">genitive</interp>
  <interp  xml:id="fn">finite clause</interp>
  <interp  xml:id="pr">prepositional</interp>
  <interp  xml:id="v1">coordinate start</interp>
</interpGrp>
```

Alternatively, a ‘stand-off’ representation for these analyses might be created using the <linkGrp> element. In this case, each linguistic segment to be annotated must be supplied with its own *xml:id* attribute:

```
<s>
  <w  xml:id="word-1">The</w>
  <w  xml:id="word-2">victim</w>
  <w  xml:id="word-3">'s</w>
  <w  xml:id="word-4">friends</w>
  <w  xml:id="word-5">told</w>
  <w  xml:id="word-6">police</w>
  <w  xml:id="word-7">that</w>
  <w  xml:id="word-8">Kruger</w>
  <w  xml:id="word-9">drove</w>
  <w  xml:id="word10">into</w>
  <w  xml:id="word11">the</w>
  <w  xml:id="word12">quarry</w>
  <w  xml:id="word13">and</w>
  <w  xml:id="word14">never</w>
  <w  xml:id="word15">surfaced</w>
</s>
```

Each segment-interpretation pair may now be represented by means of a <link> element inside an appropriate <linkGrp> element:

```
<linkGrp  type="POS-annotation">
  <link  target="#word-1 #AT0"/>
  <link  target="#word-2 #NN1"/>
  <link  target="#word-3 #POS"/>
  <link  target="#word-4 #NN2"/>
  <link  target="#word-5 #VVD"/>
  <link  target="#word-6 #NN2"/>
<!--... -->
</linkGrp>
```

Each linguistic segment so far discussed has been well-behaved with respect to the basic document hierarchy, having only a single parent. Moreover, the segmentation has been complete, in that each part of the text is accounted for by some segment at each level of analysis, without discontinuities or overlap. This state of affairs does not of course apply in all types of analysis, and these Guidelines provide a number of mechanisms to support the representation of discontinuities or multiple analyses. A brief overview of these facilities is provided in chapter *21. Non-hierarchical Structures*; also see *17. Linking, Segmentation, and Alignment.* These mechanisms all depend to a greater or lesser degree on the use of pointing elements of various kinds.

### 18.4.2 Lightweight Linguistic Annotation

While these Guidelines offer a variety of means to add linguistic information to textual units and much of that has been presented above, two kinds of use cases and two groups of users call for a dedicated set of specialized attributes to carry linguistic information. One relevant use case is where basic linguistic information gets added to an existing resource, in which generic attributes such as *type* or *ana* have already been used to encode other categorizations and analyses. The other group of users and use cases involves corpus linguists and resources built from scratch as lightly annotated language corpora. In the latter kind of projects, energy and person-hours are not devoted to careful literary analysis and hand-encoding of the relevant phenomena, but rather to the analysis of the completed resources, and therefore the phase of resource-building must be quick and relatively effortless, requiring minimal structural markup, well-established containers for grammatical information, and a standardized way of filling them in.

The aims defined above can be realized by means of lightweight linguistic annotation using attributes that belong to the `att.linguistic` class:

`att.linguistic` provides a set of attributes concerning linguistic features of tokens, for usage within token-level elements, specifically <w> and <pc> in the analysis module.

`@lemma` provides a lemma (base form) for the word, typically uninflected and serving both as an identifier (e.g. in dictionary contexts, as a headword), and as a basis for potential inflections.

`@pos` (part of speech) indicates the part of speech assigned to a token (i.e. information on whether it is a noun, adjective, or verb), usually according to some official reference vocabulary (e.g. for German: STTS, for English: CLAWS, for Polish: NKJP, etc.).

`@msd` (morphosyntactic description) supplies morphosyntactic information for a token, usually according to some official reference vocabulary (e.g. for German: STTS-large tagset; for a feature description system designed as (pragmatically) universal, see Universal Features).

`@join` when present, provides information on whether the token in question is adjacent to another, and if so, on which side.

The essence of lightweight linguistic annotation is that the basic grammatical information is encapsulated at the word level, together with the orthographic shape of the word. This has clear advantages for automatic processing but, on the other hand, this form of data encapsulation also imposes restrictions on the extent of information that can be encoded, essentially limiting it to a single tokenization and lemmatization schema, a single tagset, and a subset of the possible analyses (out from potentially many guesses at the part-of-speech or morphosyntactic descriptions, single values have to fit into the existing attributes). Another important principle that this kind of annotation is sensitive to is the need for (near) homomorphism between the assumed tokenization (division of the text stream into minimal units) and the division into minimal syntactic units (*word forms*, in the terminology of ISO Morpho-Syntactic Framework, ISO 24611[114]), because it is the former that results from the process of tokenization, but the latter that can be lemmatized and meaningfully described by means of grammatical features. Where tokens are only minimally mismatched with word forms, various repair strategies can be used (e.g., recursing <w> to capture multi-token compounds or using `att.fragmentable` to point at disjoint tokens). Beyond that, more robust TEI mechanisms, based on standoff principles and feature structures, should replace lightweight annotation.

The basic grammatical information encoded by means of `att.linguistic` is sufficient for the purpose of enhancing queries or improving the analysis of search results by, for example, making it possible to distinguish between the noun *cut* and the identically spelled verb *cut* in English, and further between e.g. the present-tense form of *cut* and its past-tense or past-participial forms. For the former contrast, the part-of-speech (*pos*) attribute should be used, whereas the latter may use *pos* and/or *msd* attributes, depending on the annotation vocabulary adopted for the project in question. The various grammatical realizations of a single 'dictionary word' can be captured by means of the attribute *lemma*, which provides a common label for them. For example, English verbs are typically lemmatized as the base form (also called *bare infinitive*), so the value of *lemma* for the verbal forms *write*, *writes*, *wrote*, *written*, and *writing* is typically write.

Together with the span-delimiting elements mentioned in this section, such as <s>, <cl>, or <phr>, lightweight grammatical annotation may be used to build basic syntactic constituency structures, where hierarchical information is expressed through span containment rather than by relations among tree nodes. This is however the limit of this kind of annotation: for the purpose of describing true constituency or dependency syntactic structures, one needs to turn to more robust mechanisms offered by the TEI, which may involve graph description (see chapter *20. Graphs, Networks, and Trees*) or standoff techniques (see section *17.9. Stand-off Markup*), and where grammatical labels may need to be annotated by means of feature structures (see chapter *19. Feature Structures*).

Some of the above-mentioned robust methods will also prove handy in cases where more than one tagset (label inventory) is used to label the words, or where automatic morphological analysis yields multiple possibilities (for example, the form *cutting* is morphologically ambiguous between verbal, adjectival, and nominal) and needs to be followed by (often also automatic) disambiguation in morphosyntactic contexts, with varying probabilities that may also need to be recorded together with their corresponding part-of-speech and morphosyntactic values.

It should be borne in mind that tokenization, lemmatization, part-of-speech identification, and morphosyntactic labelling, especially when performed automatically, should in most cases be seen as involving pragmatic decisions, dictated by concrete practical goals, economy of description, or the demands of particular analytic and/or visualization tools. It comes therefore as no surprise that numerous alternative (and often conflicting) lemmatization strategies and tagsets exist, in use by various communities and various tools, and that they change with time (a case in point is the CLAWS tagset for English, with several versions that merge the part-of-speech and morphosyntactic

[114] All definitions contained within ISO standards can be accessed at the ISO Online Browsing Platform. For ISO MAF, see `https://www.iso.org/obp/ui#iso:std:iso:24611:ed-1:v1:en`.

information to various degrees). [115] The nature and description of these systems is outside the scope of these Guidelines, but it has to be stressed that all the strategies adopted for linguistic annotation, even at the lightweight level of complexity, *must* be documented in the header of the given electronic resource, not only for the purpose of guaranteeing successful data interpretation and exchange, but also for the sake of sustainability of the results of the given project.

The last of the att.linguistic attributes, *join*, has the most text-technological flavour. It can be used to amend the loss of whitespace-related information in non-inline markup.

Compare the following two listings. The first difference between them is in the tagset used (CLAWS-5 vs. CLAWS-7) and only serves to exemplify the need to document the choice of descriptive vocabulary in the header, lest the encoded information is unreadable or confusing. The second difference is the difference in the treatment of inter-token whitespace, and it is here that the *join* attribute proves indispensable.

The first example listing uses CLAWS-5 and inline annotation, where whitespace serves as part of the markup:

```
<s><w pos="AT0">The</w> <w pos="NN1">victim</w><w pos="POS">'s</w> <w pos="NN2">friends</w>
   <w pos="VVD">told</w> <w pos="NN2">police</w> <w pos="CJT">that</w> <w pos="NP0">Kruger</w>
   <w pos="VVD">drove</w> <w pos="PRP">into</w> <w pos="AT0">the</w> <w pos="NN1">quarry</w>
   <w pos="CJC">and</w> <w pos="AV0">never</w> <w pos="SENT">surfaced</w><pc pos="PUN">.</pc></
```

In the second example, the attribute *join* is the only way to encode whether two tokens are adjacent or not:

```
<s>
  <w pos="AT">The</w>
  <w pos="NN1">victim</w>
  <w pos="GE" join="left">'s</w>
  <w pos="NN2">friends</w>
  <w pos="VVD">told</w>
  <w pos="NN2">police</w>
  <w pos="CST">that</w>
  <w pos="NP1">Kruger</w>
  <w pos="VVD">drove</w>
  <w pos="II">into</w>
  <w pos="AT">the</w>
  <w pos="NN1">quarry</w>
  <w pos="CC">and</w>
  <w pos="RR">never</w>
  <w pos="VVD">surfaced</w>
  <pc pos="." join="left">.</pc>
</s>
```

Note that projects will need to decide whether they want to redundantly encode full information on the adjacency of each token (in which case, the above listing should also have *join* with the value right on the tokens *victim* and *surfaced*, or whether information on a single direction of adjacency is enough. Strategies vary, and it is important to document them in the TEI header.

The following example shows a German sentence *Wir fahren in den Urlaub.* ('We're going on vacation.') annotated with all the attributes discussed above.[116]

[115] Given that the English language has relatively poor inflectional morphology, the decision to merge part-of-speech symbols with morphosyntactic features (as in, e.g., CLAWS-7, where the value PPHO1 signals the 3rd person singular objective personal pronoun) is fully justified as the most economical approach. For languages with more robust inflection, the *pos* and *msd* attributes will either be used separately, or the part-of-speech information will be merged into the morphosyntactic description.

[116] The annotation values have been adapted from the CLARIN Weblicht service, where e.g. the full morphosyntactic description of the first item reads: `[cat pronoun, personal true, substituting true, person 1, case nominative, number plural]`, and has been mapped from a sequence of attribute-value pairs suitable for feature structure notation, into a compressed form that fits inside a single attribute value.

```
<s>
  <w  pos="PPER"  lemma="wir"
    msd="pers:subst:p1:nom:pl">Wir</w>
  <w  pos="VVFIN"  lemma="fahren"
    msd="p1:pl:pres:ind">fahren</w>
  <w  pos="APPR"  lemma="in"  msd="--">in</w>
  <w  pos="ART"  lemma="d"
    msd="def:acc:sg:masc">den</w>
  <w  pos="NN"  lemma="Urlaub"
    msd="acc:sg:masc">Urlaub</w>
  <pc  pos="$."  lemma="."  msd="--"  join="left">.</pc>
</s>
```

The final examples lay out a strategy for dealing with e.g. historical corpora where it is on the one hand important to maintain a steady stream of token-level elements (<w> and <pc>) for efficient processing, but, on the other hand, it is also important to either record the original spelling (when the corpus text is normalized) or to record the normalized variants (when the element content of the corpus preserves the original spelling). The attribute class `att.lexicographic.normalized` can be used for that purpose:

`att.lexicographic.normalized` provides attributes for usage within word-level elements in the analysis module and within lexicographic microstructure in the dictionaries module.

- `@norm` (normalized) provides the normalized/standardized form of information present in the source text in a non-normalized form.
- `@orig` (original) gives the original string or is the empty string when the element does not appear in the source text.

The first fragment below comes from ”Gottfried, Newe Welt Vnd Americanische Historien. Frankfurt/M., 1631” encoded in the Deutsches Textarchiv and records normalized forms in the *norm* attribute.

```
<w  norm="unvermutete">vnuermuthete</w>
<w  norm="Freundschaft">Freundſchafft</w>
<w  norm="angeboten">angebotten</w>
```

The following example comes from the EarlyPrint project and uses the attribute *orig* to record the original spelling (note that the *xml:id* attributes have been removed for the sake of readability).

```
<w  lemma="he"  pos="pns">he</w>
<w  lemma="have"  pos="vvz">hath</w>
<w  lemma="bring"  pos="vvn">brought</w>
<w  lemma="forth"  pos="av"  orig="sorth">forth</w>
```

### 18.4.3 Spoken Text

The mechanisms proposed in this chapter may also be used to encode analyses of an entirely different kind, for example discourse function. Here is an application of the span technique to record details of a sales transaction in a spoken text.

```
<u  xml:id="u1">Can I have ten oranges and a kilo of bananas please?</u>
<u  xml:id="u2">Yes, anything else?</u>
<u  xml:id="u3">No thanks.</u>
<u  xml:id="u4">That'll be dollar forty.</u>
<u  xml:id="u5">Two dollars</u>
<u  xml:id="u6">Sixty, eighty, two dollars. Thank you.</u>
```

```
<spanGrp  type="transactions">
  <span  from="#u1">sale request</span>
  <span  from="#u2"  to="#u3">sale compliance</span>
  <span  from="#u4">sale</span>
  <span  from="#u5">purchase</span>
  <span  from="#u6">purchase closure</span>
</spanGrp>
```

For further discussion of the <u> (utterance) element and other elements recommended for transcriptions of spoken language, see chapter *8. Transcriptions of Speech*.

## 18.5 Module for Analysis and Interpretation

The module described in this chapter makes available the following components:

**Module analysis: Simple analytic mechanisms**

- *Elements defined*: c cl interp interpGrp m pc phr s span spanGrp w
- *Classes defined*: att.global.analytic att.lexicographic.normalized att.linguistic

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 19 Feature Structures

A *feature structure* is a general purpose data structure which identifies and groups together individual *features*, each of which associates a name with one or more values. Because of the generality of feature structures, they can be used to represent many different kinds of information, but they are of particular usefulness in the representation of linguistic analyses, especially where such analyses are partial, or *underspecified*. Feature structures represent the interrelations among various pieces of information, and their instantiation in markup provides a *metalanguage* for the generic representation of analyses and interpretations. Moreover, this instantiation allows feature values to be of specific *types*, and for restrictions to be placed on the values for particular features, by means of *feature system declarations*.[117]

## 19.1 Organization of this Chapter

This chapter is organized as follows. Following this introduction, section *19.2. Elementary Feature Structures and the Binary Feature Value* introduces the elements <fs> and <f>, used to represent feature structures and features respectively, together with the elementary *binary* feature value. Section *19.3. Other Atomic Feature Values* introduces elements for representing other kinds of atomic feature values such as *symbolic*, *numeric*, and *string* values. Section *19.4. Feature Libraries and Feature-Value Libraries* introduces the notion of predefined *libraries* or groups of features or feature values along with methods for referencing their components. Section *19.5. Feature Structures as Complex Feature Values* introduces complex values, in particular feature-structures as values, thus enabling feature structures to be recursively defined. Section *19.7. Collections as Complex Feature Values* discusses other complex values, in particular values which are collections, organized as *set*s, *bag*s, and *list*s. Section *19.8. Feature Value Expressions* discusses how the operations of alternation, negation, and collection of feature values may be represented. Section *19.9. Default Values* discusses ways of representing underspecified, default, or uncertain values. Section *19.10. Linking Text and Analysis* discusses how analyses may be linked to other parts of an encoded text. Section *19.11. Feature System Declaration* describes the *feature system declaration*, a construct which provides for the validation of typed feature structures. Formal definitions for all the elements introduced in this chapter are provided in section *19.12. Formal Definition and Implementation*.

## 19.2 Elementary Feature Structures and the Binary Feature Value

The fundamental elements used to represent a feature structure analysis are <f> (for *feature*), which represents a feature-value pair, and <fs> (for *feature structure*), which represents a structure made up of such feature-value pairs. The <fs> element has an optional *type* attribute which may be used to represent typed feature structures, and may contain any number of <f> elements. An <f> element has a required *name* attribute and an associated *value*. The value may be simple: that is, a single binary, numeric, symbolic (i.e. taken from a restricted set of legal values), or string value, or a collection of such values, organized in various ways, for example, as a list; or it may be complex, that is, it may itself be a feature structure, thus providing a degree of recursion. Values may be under-specified or defaulted in various ways. These possibilities are all described in more detail in this and the following sections.

Feature and feature-value representations (including feature structure representations) may be embedded directly at any point in an XML document, or they may be collected together in special-purpose feature or feature-value *libraries*. The components of such libraries may then be referenced from other feature or feature-value representations, using the *feats* or *fVal* attribute as appropriate.

We begin by considering the simple case of a feature structure which contains binary-valued features only. The following three XML elements are needed to represent such a feature structure:

`<fs>` (feature structure) represents a *feature structure*, that is, a collection of feature-value pairs organized as a structural unit.

`@type` specifies the type of the feature structure.

`@feats` (features) references the feature-value specifications making up this feature structure.

`<f>` (feature) represents a *feature value specification*, that is, the association of a name with a value of any of several different types.

[117] The recommendations of this chapter have been adopted as ISO Standard 24610-1 *Language Resource Management — Feature Structures — Part One: Feature Structure Representation*

@*name* a single word which follows the rules defining a legal XML name (see `https://www.w3.org/TR/REC-xml/#dt-name`), providing a name for the feature.

@*fVal* (feature value) references any element which can be used to represent the value of a feature.

`<binary>` (binary value) represents the value part of a feature-value specification which can contain either of exactly two possible values.

The attributes *feats* and the *fVal* are not discussed in this section: they provide an alternative way of indicating the content of an element, as further discussed in section *19.4. Feature Libraries and Feature-Value Libraries*.

An <fs> element containing <f> elements with binary values can be straightforwardly used to encode the *matrices* of feature-value specifications for phonetic segments, such as the following for the English segment [s].

```
  +--- ---+ | consonantal + | | vocalic - | | voiced - | | anterior + | | coronal + |
| continuant + | | strident + | +--- ---+
```

This representation may be encoded in XML as follows:

```
<fs  type="phonological_segments">
  <f  name="consonantal">
    <binary  value="true"/>
  </f>
  <f  name="vocalic">
    <binary  value="false"/>
  </f>
  <f  name="voiced">
    <binary  value="false"/>
  </f>
  <f  name="anterior">
    <binary  value="true"/>
  </f>
  <f  name="coronal">
    <binary  value="true"/>
  </f>
  <f  name="continuant">
    <binary  value="true"/>
  </f>
  <f  name="strident">
    <binary  value="true"/>
  </f>
</fs>
```

Note that <fs> elements may have an optional *type* attribute to indicate the kind of feature structure in question, whereas <f> elements must have a *name* attribute to indicate the name of the feature. Feature structures need not be typed, but features must be named. Similarly, the <fs> element may be empty, but the <f> element must specify its value either directly as content, by means of the *fVal* attribute, or implicitly by reference to a feature system declaration.

The restriction of specific features to specific types of values (e.g. the restriction of the feature *strident* to a binary value) requires additional validation, as does any restriction on the features available within a feature structure of a particular type (e.g. whether a feature structure of type *phonological segment* necessarily contains a feature *voiced*). Such validation may be carried out at the document level, using special purpose processing, at the schema level using additional validation rules, or at the declarative level, using an additional mechanism such as the *feature-system declaration* discussed in *19.11. Feature System Declaration*.

Although we have used the term *binary* for this kind of value, and its representation in XML uses values such as `true` and `false` (or, equivalently, `1` and `0`), it should be noted that such values are not restricted to propositional assertions. As this example shows, this kind of value is intended for use with any binary-valued feature.

## 19.3 Other Atomic Feature Values

Features may take other kinds of atomic value. In this section, we define elements which may be used to represent: *symbolic values*, *numeric values*, and *string values*. The module defined by this chapter allows for the specification of additional datatypes if necessary, by extending the underlying class `model.featureVal.single`. If this is done, it is recommended that only the basic W3C datatypes should be used; more complex datatyping should be represented as feature structures.

`<symbol>` (symbolic value) represents the value part of a feature-value specification which contains one of a finite list of symbols.

`@value` supplies a symbolic value for the feature, one of a finite list that may be specified in a feature declaration.

`<numeric>` (numeric value) represents the value part of a feature-value specification which contains a numeric value or range.

`<string>` (string value) represents the value part of a feature-value specification which contains a string.

The <symbol> element is used for the value of a feature when that feature can have any of a small, finite set of possible values, representable as character strings. For example, the following might be used to represent the claim that the Latin noun form *mensas* (tables) has accusative case, feminine gender, and plural number:

```
<fs>
  <f  name="case">
    <symbol  value="accusative"/>
  </f>
  <f  name="gender">
    <symbol  value="feminine"/>
  </f>
  <f  name="number">
    <symbol  value="plural"/>
  </f>
</fs>
```

More formally, this representation shows a structure in which three features (*case*, *gender*, and *number*) are used to define morpho-syntactic properties of a word. Each of these features can take one of a small number of values (for example, case can be `nominative`, `genitive`, `dative`, `accusative`, etc.) and it is therefore appropriate to represent the values taken in this instance as <symbol> elements. Note that, instead of using a symbolic value for grammatical number, one could have named the feature *singular* or *plural* and given it an appropriate binary value, as in the following example:

```
<fs>
  <f  name="case">
    <symbol  value="accusative"/>
  </f>
  <f  name="gender">
    <symbol  value="feminine"/>
  </f>
  <f  name="singular">
    <binary  value="false"/>
  </f>
</fs>
```

Whether one uses a binary or symbolic value in situations like this is largely a matter of taste.

The <string> element is used for the value of a feature when that value is a string drawn from a very large or potentially unbounded set of possible strings of characters, so that it would be impractical or impossible to use the <symbol> element. The string value is expressed as the content of the <string> element, rather than as an attribute value. For example, one might encode a street address as follows:

```
<fs>
  <f  name="address">
    <string>3418 East Third Street</string>
  </f>
</fs>
```

The <numeric> element is used when the value of a feature is a numeric value, or a range of such values. For example, one might wish to regard the house number and the street name as different features, using an encoding like the following:

```
<fs>
  <f  name="houseNumber">
    <numeric  value="3418"/>
  </f>
  <f  name="streetName">
    <string>East Third Street</string>
  </f>
</fs>
```

If the numeric value to be represented falls within a specific range (for example an address that spans several numbers), the *max* attribute may be used to supply an upper limit:

```
<fs>
  <f  name="houseNumber">
    <numeric  value="3418"  max="3440"/>
  </f>
  <f  name="streetName">
    <string>East Third Street</string>
  </f>
</fs>
```

It is also possible to specify that the numeric value (or values) represented should (or should not) be truncated. For example, assuming that the daily rainfall in mm is a feature of interest for some address, one might represent this by an encoding like the following:

```
<fs>
  <f  name="dailyRainFall">
    <numeric  value="0.0"  max="1.3"
      trunc="false"/>
  </f>
</fs>
```

This represents any of the infinite number of numeric values falling between 0 and 1.3; by contrast

```
<fs>
  <f  name="dailyRainFall">
    <numeric  value="0.0"  max="1.3"
      trunc="true"/>
  </f>
</fs>
```

represents only two possible values: 0 and 1.

Some communities of practice, notably those with a strong computer-science bias, prefer to dissociate the information on the value of the given feature from the specification of the data type that this value represents. In such cases, feature values can be provided directly as textual content of <f>, with the assumption that the data type is specified by the schema. The following is an example taken from ISO 24612, presenting the symbolic values for Active Voice and Simple Present Tense in the untyped form:

```
<fs>
  <f  name="voice">active</f>
  <f  name="tense">SimPre</f>
</fs>
```

As noted above, additional processing is necessary to ensure that appropriate values are supplied for particular features, for example to ensure that the feature `singular` is not given a value such as <symbol value="feminine"/>. There are two ways of attempting to ensure that only certain combinations of feature names and values are used. First, if the total number of legal combinations is relatively small, one can predefine all of them in a construct known as a *feature library*, and then reference the combination required using the *feats* attribute in the enclosing <fs> element, rather than give it explicitly. This method is suitable in the situation described above, since it requires specifying a total of only ten (5 + 3 + 2) combinations of features and values. Similarly, to ensure that only feature structures containing valid combinations of feature values are used, one can put definitions for all valid feature structures inside a *feature value library* (so called, since a feature structure may be the value of a feature). A total of 30 feature structures (5 × 3 × 2) is required to enumerate all the possible combinations of individual case, gender and number values in the preceding illustration. We discuss the use of such libraries and their representation in XML further in section *19.4. Feature Libraries and Feature-Value Libraries* below.

However, the most general method of attempting to ensure that only legal combinations of feature names and values are used is to provide a *feature-system declaration* discussed in *19.11. Feature System Declaration.*

Whether at the level of feature-system declarations, feature- and feature-value libraries, or individual features, it is possible to align both feature names and their values with standardized external data category repositories. [118] In the following example, both the feature part_of_speech and its value NN (standing for 'common noun') are aligned with the respective definitions provided by the CLARIN Concept Registry (CCR).

```
<fs>
  <f  name="part_of_speech"
    datcat="http://hdl.handle.net/11459/CCR_C-396_5a972b93-2294-ab5c-a541-7c344c5f26c3">
    <symbol  valueDatcat="http://hdl.handle.net/11459/CCR_C-1256_7ec6083c-23d4-224d-6f94-eecbe6861
      value="NN"/>
  </f>
<!-- ... -->
</fs>
```

Since the above representation takes up a lot of space and quickly becomes redundant and error-prone, it is possible to delegate the task of aligning with external repositories to elements such as <fLib>, <fvLib>, <fDecl>, or <fsDecl> to reduce the feature representation at hand and to increase its readability at the same time, as shown in the example below.

```
<fs>
<!-- ... -->
  <f  name="POS"  fVal="#common_noun"/>
<!-- ... -->
</fs>
```

[118] See section *10.5.2. Lexical View* for more discussion of the need and rationale for aligning the content of grammatical (and other) descriptions with standardized external taxonomies.

The value common_noun should best be listed (as an *xml:id*) either in a library of feature values (<fvLib>, see the following section) or in a <taxonomy> element.

## 19.4 Feature Libraries and Feature-Value Libraries

As the examples in the preceding section suggest, the direct encoding of feature structures can be verbose. Moreover, it is often the case that particular feature-value combinations, or feature structures composed of them, are re-used in different analyses. To reduce the size and complexity of the task of encoding feature structures, one may use the *feats* attribute of the <fs> element to point to one or more of the feature-value specifications for that element. This indirect method of encoding feature structures presumes that the <f> elements are assigned unique *xml:id* values, and are collected together in <fLib> elements (*feature libraries*). In the same way, feature values of whatever type can be collected together in <fvLib> elements (*feature-value libraries*). If a feature has as its value a feature structure or other value which is predefined in this way, the *fVal* attribute may be used to point to it, as discussed in the next section. The following elements are used for representing feature libraries and feature-value libraries:

`<fLib>` (feature library) assembles a library of <f> (feature) elements.

`<fvLib>` (feature-value library) assembles a library of reusable feature value elements (including complete feature structures).

For example, suppose a feature library for phonological feature specifications is set up as follows.

```
<fLib n="phonological features">
  <f xml:id="CNS1" name="consonantal">
    <binary value="true"/>
  </f>
  <f xml:id="CNS0" name="consonantal">
    <binary value="false"/>
  </f>
  <f xml:id="VOC1" name="vocalic">
    <binary value="true"/>
  </f>
  <f xml:id="VOC0" name="vocalic">
    <binary value="false"/>
  </f>
  <f xml:id="VOI1" name="voiced">
    <binary value="true"/>
  </f>
  <f xml:id="VOI0" name="voiced">
    <binary value="false"/>
  </f>
  <f xml:id="ANT1" name="anterior">
    <binary value="true"/>
  </f>
  <f xml:id="ANT0" name="anterior">
    <binary value="false"/>
  </f>
  <f xml:id="COR1" name="coronal">
    <binary value="true"/>
  </f>
  <f xml:id="COR0" name="coronal">
    <binary value="false"/>
  </f>
  <f xml:id="CNT1" name="continuant">
    <binary value="true"/>
  </f>
  <f xml:id="CNT0" name="continuant">
    <binary value="false"/>
```

```
  </f>
  <f  xml:id="STR1"  name="strident">
    <binary  value="true"/>
  </f>
  <f  xml:id="STR0"  name="strident">
    <binary  value="false"/>
  </f>
<!-- ... -->
</fLib>
```

Then the feature structures that represent the analysis of the phonological segments (phonemes) /t/, /d/, /s/, and /z/ may be defined as follows.

```
<fs  feats="#CNS1 #VOC0 #VOI0 #ANT1 #COR1 #CNT0 #STR0"/>
<fs  feats="#CNS1 #VOC0 #VOI1 #ANT1 #COR1 #CNT0 #STR0"/>
<fs  feats="#CNS1 #VOC0 #VOI0 #ANT1 #COR1 #CNT1 #STR1"/>
<fs  feats="#CNS1 #VOC0 #VOI1 #ANT1 #COR1 #CNT1 #STR1"/>
```

The preceding are but four of the 128 logically possible fully specified phonological segments using the seven binary features listed in the feature library. Presumably not all combinations of features correspond to phonological segments (there are no strident vowels, for example). The legal combinations, however, can be collected together, each one represented as an identifiable <fs> element within a *feature-value library*, as in the following example:

```
<fvLib  xml:id="fsl1"
  n="phonological segment definitions">
<!-- ... -->
  <fs  xml:id="T.DF"
    feats="#CNS1 #VOC0 #VOI0 #ANT1 #COR1 #CNT0 #STR0"/>
  <fs  xml:id="D.DF"
    feats="#CNS1 #VOC0 #VOI1 #ANT1 #COR1 #CNT0 #STR0"/>
  <fs  xml:id="S.DF"
    feats="#CNS1 #VOC0 #VOI0 #ANT1 #COR1 #CNT1 #STR1"/>
  <fs  xml:id="Z.DF"
    feats="#CNS1 #VOC0 #VOI1 #ANT1 #COR1 #CNT1 #STR1"/>
<!-- ... -->
</fvLib>
```

Once defined, these feature structure values can also be reused. Other <f> elements may invoke them by reference, using the *fVal* attribute; for example, one might use them in a feature value pair such as:

```
<f  name="dental-fricative"  fVal="#T.DF"/>
```

rather than expanding the hierarchy of the component phonological features explicitly.

The feature structure that concludes section *19.3. Other Atomic Feature Values* above, identifying the value of some part of speech to be a common noun, may be used in tandem with a feature-value library, which offers a way to encode a grammatical *tagset*, in this case containing labels for parts of speech:

```
<fvLib  n="POS values">
  <symbol  xml:id="common_noun"  value="NN"
    datcat="http://hdl.handle.net/11459/CCR_C-396_5a972b93-2294-ab5c-a541-7c344c5f26c3"/>
  <symbol  xml:id="proper_noun"  value="NP"
```

```
   datcat="http://hdl.handle.net/11459/CCR_C-1371_fbebd9ec-a7f4-9a36-d6e9-88ee16b944ae"/>
<!-- ... -->
</fvLib>
```

Such a feature-value library combines the standard short symbolic label for a part of speech (e.g., NN) with a mnemonic identifier that can be referenced by means of *fVal*, and with a persistent identifier, maintained in a public reference taxonomy repository together with the basic definition of the given concept.

Feature structures stored in the way presented in this section may also be associated with the text which they are intended to annotate, either by a link from the text (for example, using the TEI global *ana* attribute), or by means of stand-off annotation techniques (for example, using the TEI <link> element): see further section *19.10. Linking Text and Analysis* below.

Note that when features or feature structures are linked to in this way, the result is effectively a copy of the item linked to into the place from which it is linked. This form of linking should be distinguished from the phenomenon of *structure-sharing*, where it is desired to indicate that some part of an annotation structure appears simultaneously in two or more places within the structure. This kind of annotation should be represented using the <vLabel> element, as discussed in *19.6. Re-entrant Feature Structures* below.

## 19.5 Feature Structures as Complex Feature Values

Features may have complex values as well as atomic ones; the simplest such complex value is represented by supplying an <fs> element as the content of an <f> element, or (equivalently) by supplying the identifier of an <fs> element as the value for the *fVal* attribute on the <f> element. Structures may be nested as deeply as appropriate, using this mechanism. For example, an <fs> element may contain or point to an <f> element, which may contain or point to an <fs> element, which may contain or point to an <f> element, and so on.

To illustrate the use of complex values, consider the following simple model of a word, as a structure combining surface form information, a syntactic category, and semantic information. Each word analysis is represented as a <fs type='word'> element, containing three features named `surface`, `syntax`, and `semantics`. The first of these has an atomic string value, but the other two have complex values, represented as nested feature structures of types `category` and `act` respectively:

```
<fs type="word">
  <f name="surface">
    <string>love</string>
  </f>
  <f name="syntax">
    <fs type="category">
      <f name="pos">
        <symbol value="verb"/>
      </f>
      <f name="val">
        <symbol value="transitive"/>
      </f>
    </fs>
  </f>
  <f name="semantics">
    <fs type="act">
      <f name="rel">
        <symbol value="LOVE"/>
      </f>
    </fs>
  </f>
</fs>
```

This analysis does not tell us much about the meaning of the symbols `verb` or `transitive`. It might be preferable to replace these atomic feature values by feature structures. Suppose therefore that we maintain a feature-value library for each of the major syntactic categories (N, V, ADJ, PREP):

```
<fvLib  n="Major category definitions">
<!-- ... -->
  <fs  xml:id="N"  type="noun">
<!-- noun features defined here -->
  </fs>
  <fs  xml:id="V"  type="verb">
<!-- verb features defined here -->
  </fs>
</fvLib>
```

This library allows us to use shortcut codes (`N`, `V`, etc.) to reference a complete definition for the corresponding feature structure. Each definition may be explicitly contained within the <fs> element, as a number of <f> elements. Alternatively, the relevant features may be referenced by their identifiers, supplied as the value of the *feats* attribute, as in these examples:

```
<!-- ... -->
<fs xml:id="ADJ" type="adjective" feats="#F1 #F2"/>
<fs xml:id="PREP" type="preposition" feats="#F1 #F3"/>
<!-- ... -->
```

This ability to re-use feature definitions within multiple feature structure definitions is an essential simplification in any realistic example. In this case, we assume the existence of a feature library containing specifications for the basic feature categories like the following:

```
<fLib  n="categorial features">
  <f  xml:id="NN-1"  name="nominal">
    <binary  value="true"/>
  </f>
  <f  xml:id="NN-0"  name="nominal">
    <binary  value="false"/>
  </f>
  <f  xml:id="VV-1"  name="verbal">
    <binary  value="true"/>
  </f>
  <f  xml:id="VV-0"  name="verbal">
    <binary  value="false"/>
  </f>
<!-- ... -->
</fLib>
```

With such libraries in place, and assuming the availability of similarly predefined feature structures for transitivity and semantics, the preceding example could be considerably simplified:

```
<fs  type="word">
  <f  name="surface">
    <string>love</string>
  </f>
```

```
  <f  name="syntax">
    <fs  type="category">
      <f  name="pos"  fVal="#V"/>
      <f  name="val"  fVal="#TRNS"/>
    </fs>
  </f>
  <f  name="semantics">
    <fs  type="act">
      <f  name="rel"  fVal="#LOVE"/>
    </fs>
  </f>
</fs>
```

Although in principle the *fVal* attribute could point to any kind of feature value, its use is not recommended for simple atomic values.

## 19.6 Re-entrant Feature Structures

Sometimes the same feature value is required at multiple places within a feature structure, in particular where the value is only partially specified at one or more places. The <vLabel> element is provided as a means of labelling each such re-entrancy point:

`<vLabel>` (value label) represents the value part of a feature-value specification which appears at more than one point in a feature structure.

For example, suppose one wishes to represent noun-verb agreement as a single feature structure. Within the representation, the feature indicating (say) number appears more than once. To represent the fact that each occurrence is another appearance of the same feature (rather than a copy) one could use an encoding like the following:

```
<fs  xml:id="NVA">
  <f  name="nominal">
    <fs>
      <f  name="nm-num">
        <vLabel  name="L1">
          <symbol  value="singular"/>
        </vLabel>
      </f>
<!-- other nominal features -->
    </fs>
  </f>
  <f  name="verbal">
    <fs>
      <f  name="vb-num">
        <vLabel  name="L1"/>
      </f>
    </fs>
<!-- other verbal features -->
  </f>
</fs>
```

In the above encoding, the features named `vb-num` and `nm-num` exhibit structure sharing. Their values, given as `vLabel` elements, are understood to be references to the same point in the feature structure, which is labelled by their *name* attribute.

The scope of the names used to label re-entrancy points is that of the outermost <fs> element in which they appear. When a feature structure is imported from a feature value library, or referenced from elsewhere (for example by using the *fVal* attribute) the names of any sharing points it may contain are implicitly prefixed by the identifier used for

the imported feature structure, to avoid name clashes. Thus, if some other feature structure were to reference the <fs> element given in the example above, for example in this way:

```
<f  name="class"  fVal="#NVA"/>
```

then the labelled points in the example would be interpreted as if they had the name NVAL1.

## 19.7 Collections as Complex Feature Values

Complex feature values need not always be represented as feature structures. Multiple values may also be organized as sets, bags or multisets, or lists of atomic values of any type. The <vColl> element is provided to represent such cases:

**<vColl>** (collection of values) represents the value part of a feature-value specification which contains multiple values organized as a set, bag, or list.

A feature whose value is regarded as a set, bag, or list may have any positive number of values as its content, or none at all, (thus allowing for representation of the empty set, bag, or list). The items in a list are ordered, and need not be distinct. The items in a set are not ordered, and must be distinct. The items in a bag are neither ordered nor distinct. Sets and bags are thus distinguished from lists in that the order in which the values are specified does not matter for the former, but does matter for the latter, while sets are distinguished from bags and lists in that repetitions of values do not count for the former but do count for the latter.

If no value is specified for the *org* attribute, the assumption is that the <vColl> defines a list of values. If the <vColl> element is empty, the assumption is that it represents the null list, set, or bag.

To illustrate the use of the *org* attribute, suppose that a feature structure analysis is used to represent a genealogical tree, with the information about each individual treated as a single feature structure, like this:

```
<fs  xml:id="p027"  type="person">
  <f  name="forenames">
    <vColl>
      <string>Daniel</string>
      <string>Edouard</string>
    </vColl>
  </f>
  <f  name="mother"  fVal="#p002"/>
  <f  name="father"  fVal="#p009"/>
  <f  name="birthDate">
    <fs  type="date"  feats="#y1988 #m04 #d17"/>
  </f>
  <f  name="birthPlace"  fVal="#austintx"/>
  <f  name="siblings">
    <vColl  org="set">
      <fs  copyOf="#pnb005"/>
      <fs  copyOf="#prb001"/>
    </vColl>
  </f>
</fs>
```

In this example, the <vColl> element is first used to supply a list of 'name' feature values, which together constitute the 'forenames' feature. Other features are defined by reference to values which we assume are held in some external feature value library (not shown here). For example, the <vColl> element is used a second time to indicate that the persons's siblings should be regarded as constituting a set rather than a list. Each sibling is represented by a feature structure: in this example, each feature structure is a copy of one specified in the feature value library.

If a specific feature contains only a single feature structure as its value, the component features of which are organized as a set, bag, or list, it may be more convenient to represent the value as a <vColl> rather than as an <fs>. For example, consider the following encoding of the English verb form *sinks*, which contains an *agreement* feature whose value is a feature structure which contains *person* and *number* features with symbolic values.

```
<fs type="word">
  <f name="category">
    <symbol value="verb"/>
  </f>
  <f name="tense">
    <symbol value="present"/>
  </f>
  <f name="agreement">
    <fs>
      <f name="person">
        <symbol value="third"/>
      </f>
      <f name="number">
        <symbol value="singular"/>
      </f>
    </fs>
  </f>
</fs>
```

If the names of the features contained within the *agreement* feature structure are of no particular significance, the following simpler representation may be used:

```
<fs type="word">
  <f name="category">
    <symbol value="verb"/>
  </f>
  <f name="tense">
    <symbol value="present"/>
  </f>
  <f name="agreement">
    <vColl org="set">
      <symbol value="third"/>
      <symbol value="singular"/>
    </vColl>
  </f>
</fs>
```

The <vColl> element is also useful in cases where an analysis has several components. In the following example, the French word *auxquels* has a two-part analysis, represented as a list of two values. The first specifies that the word contains a preposition; the second that it contains a masculine plural relative pronoun:

```
<fs>
  <f name="lex">
    <symbol value="auxquels"/>
  </f>
  <f name="maf">
    <vColl org="list">
      <fs>
        <f name="cat">
          <symbol value="prep"/>
        </f>
      </fs>
      <fs>
        <f name="cat">
```

```
        <symbol value="pronoun"/>
      </f>
      <f name="kind">
        <symbol value="rel"/>
      </f>
      <f name="num">
        <symbol value="pl"/>
      </f>
      <f name="gender">
        <symbol value="masc"/>
      </f>
    </fs>
  </vColl>
 </f>
</fs>
```

The set, bag, or list which has no members is known as the null (or empty) set, bag, or list. A <vColl> element with no content and with no value for its *feats* attribute is interpreted as referring to the null set, bag, or list, depending on the value of its *org* attribute.

If, for example, the individual described by the feature structure with identifier `p027` (above) had no siblings, we might specify the *siblings* feature as follows.

```
<f name="siblings">
 <vColl org="set"/>
</f>
```

A <vColl> element may also collect together one or more other <vColl> elements, if, for example one of the members of a set is itself a set, or if two lists are concatenated together. Note that such collections pay no attention to the contents of the nested <vColl> elements: if it is desired to produce the union of two sets, the <vMerge> element discussed below should be used to make a new collection from the two sets.

## 19.8 Feature Value Expressions

It is sometimes desirable to express the value of a feature as the result of an operation over some other value (for example, as 'not green', or as 'male or female', or as the concatenation of two collections). Three special purpose elements are provided to represent disjunctive alternation, negation, and collection of values:

`<vAlt>` (value alternation) represents the value part of a feature-value specification which contains a set of values, only one of which can be valid.

`<vNot>` (value negation) represents a feature value which is the negation of its content.

`<vMerge>` (merged collection of values) represents a feature value which is the result of merging together the feature values contained by its children, using the organization specified by the *org* attribute.

### 19.8.1 Alternation

The <vAlt> element can be used wherever a feature value can appear. It contains two or more feature values, any one of which is to be understood as the value required. Suppose, for example, that we are using a feature system to describe residential property, using such features as *number.of.bathrooms*. In a particular case, we might wish to represent uncertainty as to whether a house has two or three bathrooms. As we have already shown, one simple way to represent this would be with a numeric maximum:

```
<f name="number.of.bathrooms">
 <numeric value="2" max="3"/>
</f>
```

A more general way would be to represent the alternation explicitly, in this way:

```
<f  name="number.of.bathrooms">
  <vAlt>
    <numeric  value="2"/>
    <numeric  value="3"/>
  </vAlt>
</f>
```

The <vAlt> element represents alternation over feature values, not feature-value pairs. If therefore the uncertainty relates to two or more feature value specifications, each must be represented as a feature structure, since a feature structure can always appear where a value is required. For example, suppose that it is uncertain as to whether the house being described has two bathrooms or two bedrooms, a structure like the following may be used:

```
<f  name="rooms">
  <vAlt>
    <fs>
      <f  name="number.of.bathrooms">
        <numeric  value="2"/>
      </f>
    </fs>
    <fs>
      <f  name="number.of.bedrooms">
        <numeric  value="2"/>
      </f>
    </fs>
  </vAlt>
</f>
```

Note that alternation is always regarded as *exclusive*: in the case above, the implication is that having two bathrooms excludes the possibility of having two bedrooms and vice versa. If inclusive alternation is required, a <vColl> element may be included in the alternation as follows:

```
<f  name="rooms">
  <vAlt>
    <fs>
      <f  name="number.of.bathrooms">
        <numeric  value="2"/>
      </f>
    </fs>
    <fs>
      <f  name="number.of.bedrooms">
        <numeric  value="2"/>
      </f>
    </fs>
    <vColl>
      <fs>
        <f  name="number.of.bathrooms">
          <numeric  value="2"/>
        </f>
      </fs>
      <fs>
        <f  name="number.of.bedrooms">
          <numeric  value="2"/>
```

```
      </f>
     </fs>
   </vColl>
  </vAlt>
</f>
```

This analysis indicates that the property may have two bathrooms, two bedrooms, or both two bathrooms and two bedrooms.

As the previous example shows, the <vAlt> element can also be used to indicate alternations among values of features organized as sets, bags or lists. Suppose we use a feature `selling.points` to describe items that are mentioned to enhance a property's sales value, such as whether it has a pool or a good view. Now suppose for a particular listing, the selling points include an alarm system and a good view, and either a pool or a jacuzzi (but not both). This situation could be represented, using the <vAlt> element, as follows.

```
<fs  type="real_estate_listing">
  <f  name="selling.points">
    <vColl  org="set">
      <string>alarm system</string>
      <string>good view</string>
      <vAlt>
        <string>pool</string>
        <string>jacuzzi</string>
      </vAlt>
    </vColl>
  </f>
</fs>
```

Now suppose the situation is like the preceding except that one is also uncertain whether the property has an alarm system or a good view. This can be represented as follows.

```
<fs  type="real_estate_listing">
  <f  name="selling.points">
    <vColl  org="set">
      <vAlt>
        <string>alarm system</string>
        <string>good view</string>
      </vAlt>
      <vAlt>
        <string>pool</string>
        <string>jacuzzi</string>
      </vAlt>
    </vColl>
  </f>
</fs>
```

If a large number of ambiguities or uncertainties need to be represented, involving a relatively small number of features and values, it is recommended that a stand-off technique, for example using the general-purpose <alt> element discussed in section *17.8. Alternation* be used, rather than the special-purpose <vAlt> element.

### 19.8.2 Negation

The <vNot> element can be used wherever a feature value can appear. It contains any feature value and returns the complement of its contents. For example, the feature *number.of.bathrooms* in the following example has any whole numeric value other than 2:

```
<f  name="number.of.bathrooms">
  <vNot>
    <numeric  value="2"/>
  </vNot>
</f>
```

Strictly speaking, the effect of the <vNot> element is to provide the complement of the feature values it contains, rather than their negation. If a feature system declaration is available which defines the possible values for the associated feature, then it is possible to say more about the negated value. For example, suppose that the available values for the feature `case` are declared to be nominative, genitive, dative, or accusative, whether in a TEI feature system declaration or by some other means. Then the following two specifications are equivalent:

```
 (i) <f  name="case">
  <vNot>
    <symbol  value="genitive"/>
  </vNot>
</f>
(ii)
<f  name="case">
  <vAlt>
    <symbol  value="nominative"/>
    <symbol  value="dative"/>
    <symbol  value="accusative"/>
  </vAlt>
</f>
```

If however no such system declaration is available, all that one can say about a feature specified via negation is that its value is something other than the negated value.

Negation is always applied to a feature value, rather than to a feature-value pair. The negation of an atomic value is the set of all other values which are possible for the feature.

Any kind of value can be negated, including collections (represented by a <vColl> elements) or feature structures (represented by <fs> elements). The negation of any complex value is understood to be the set of values which cannot be unified with it. Thus, for example, the negation of the feature structure F is understood to be the set of feature structures which are not unifiable with F. In the absence of a constraint mechanism such as the Feature System Declaration, the negation of a collection is anything that is not unifiable with it, including collections of different types and atomic values. It will generally be more useful to require that the organization of the negated value be the same as that of the original value, for example that a negated set is understood to mean the set which is a complement of the set, but such a requirement cannot be enforced in the absence of a constraint mechanism.

### 19.8.3 Collection of Values

The <vMerge> element can be used wherever a feature value can appear. It contains two or more feature values, all of which are to be collected together. The organization of the resulting collection is specified by the value of the *org* attribute, which need not necessarily be the same as that of its constituent values if these are collections. For example, one can change a list to a set, or vice versa.

As an example, suppose that we wish to represent the range of possible values for a feature ‘genders’ used to describe some language. It would be natural to represent the possible values as a set, using the <vColl> element as in the following example:

```
<fs>
  <f  name="genders">
    <vColl  org="set">
      <symbol  value="masculine"/>
```

```
    <symbol value="feminine"/>
   </vColl>
  </f>
</fs>
```

Suppose however that we discover for some language it is necessary to add a new possible value, and to treat the value of the feature as a list rather than as a set. The <vMerge> element can be used to achieve this:

```
<fs>
  <f name="genders">
   <vMerge org="list">
    <vColl org="set">
     <symbol value="masculine"/>
     <symbol value="feminine"/>
    </vColl>
    <symbol value="neuter"/>
   </vMerge>
  </f>
</fs>
```

## 19.9 Default Values

The value of a feature may be underspecified in a number of different ways. It may be null, unknown, or uncertain with respect to a range of known possibilities, as well as being defined as a negation or an alternation. As previously noted, the specification of the range of known possibilities for a given feature is not part of the current specification: in the TEI scheme, this information is conveyed by the *feature system declaration*. Using this, or some other system, we might specify (for example) that the range of values for an element includes symbols for masculine, feminine, and neuter, and that the default value is neuter. With such definitions available to us, it becomes possible to say that some feature takes the default value, or some unspecified value from the list. The following special element is provided for this purpose:

`<default>` (default feature value) represents the value part of a feature-value specification which contains a defaulted value.

The value of an empty <f> element which also lacks an *fVal* attribute is understood to be the most general case, i.e. any of the available values. Thus, assuming the feature system defined above, the following two representations are equivalent.

```
<f name="gender"/>
<f name="gender">
  <vAlt>
   <symbol value="feminine"/>
   <symbol value="masculine"/>
   <symbol value="neuter"/>
  </vAlt>
</f>
```

If, however, the value is explicitly stated to be the default one, using the <default> element, then the following two representations are equivalent:

```
<f name="gender">
  <default/>
</f>
```

```
<f  name="gender">
  <symbol  value="neuter"/>
</f>
```

Similarly, if the value is stated to be the negation of the default, then the following two representations are equivalent:

```
<f  name="gender">
  <vNot>
    <default/>
  </vNot>
</f>
```

```
<f  name="gender">
  <vAlt>
    <symbol  value="feminine"/>
    <symbol  value="masculine"/>
  </vAlt>
</f>
```

## 19.10 Linking Text and Analysis

Text elements can be linked with feature structures using any of the linking methods discussed elsewhere in these Guidelines (see for example sections *18.2. Global Attributes for Simple Analyses* and *18.4. Linguistic Annotation*). In the simplest case, the *ana* attribute may be used to point from any element to an annotation of it, as in the following example:

```
<s  n="00741">
  <w  ana="#at0">The</w>
  <w  ana="#ajs">closest</w>
  <w  ana="#pnp">he</w>
  <w  ana="#vvd">came</w>
  <w  ana="#prp">to</w>
  <w  ana="#nn1">exercise</w>
  <w  ana="#vbd">was</w>
  <w  ana="#to0">to</w>
  <w  ana="#vvi">open</w>
  <w  ana="#crd">one</w>
  <w  ana="#nn1">eye</w>
  <phr  ana="#av0">
    <w>every</w>
    <w>so</w>
    <w>often</w>
  </phr>
  <c  ana="#pun">,</c>
  <w  ana="#cjs">if</w>
  <w  ana="#pni">someone</w>
  <w  ana="#vvd">entered</w>
  <w  ana="#at0">the</w>
  <w  ana="#nn1">room</w>
<!-- ... -->
</s>
```

The values specified for the *ana* attribute reference components of a feature-structure library, which represents all of the grammatical structures used by this encoding scheme. (For illustrative purposes, we cite here only the structures needed for the first six words of the sample sentence):

```
<fvLib  xml:id="C6"  n="Claws 6 tags">
<!-- ... -->
  <fs  xml:id="ajs"
    type="grammatical_structure"  feats="#wj #ds"/>
  <fs  xml:id="at0"
    type="grammatical_structure"  feats="#wl"/>
  <fs  xml:id="pnp"
    type="grammatical_structure"  feats="#wr #rp"/>
  <fs  xml:id="vvd"
    type="grammatical_structure"  feats="#wv #bv #fd"/>
  <fs  xml:id="prp"
    type="grammatical_structure"  feats="#wp #bp"/>
  <fs  xml:id="nnn"
    type="grammatical_structure"  feats="#wn #tc #ns"/>
<!-- ... -->
</fvLib>
```

The components of each feature structure in the library are referenced in much the same way, using the *feats* attribute to identify one or more <f> elements in the following feature library (again, only a few of the available features are quoted here):

```
<fLib>
<!-- ... -->
  <f  xml:id="fl-bv"  name="verbbase">
    <symbol  value="main"/>
  </f>
  <f  xml:id="fl-bp"  name="prepbase">
    <symbol  value="lexical"/>
  </f>
  <f  xml:id="fl-ds"  name="degree">
    <symbol  value="superlative"/>
  </f>
  <f  xml:id="fl-fd"  name="verbform">
    <symbol  value="ed"/>
  </f>
  <f  xml:id="fl-ns"  name="number">
    <symbol  value="singular"/>
  </f>
  <f  xml:id="fl-rp"  name="prontype">
    <symbol  value="personal"/>
  </f>
  <f  xml:id="fl-tc"  name="nountype">
    <symbol  value="common"/>
  </f>
  <f  xml:id="fl-wj"  name="class">
    <symbol  value="adjective"/>
  </f>
  <f  xml:id="fl-wl"  name="class">
    <symbol  value="article"/>
  </f>
  <f  xml:id="fl-wn"  name="class">
    <symbol  value="noun"/>
```

```
  </f>
  <f  xml:id="fl-wp"  name="class">
    <symbol  value="preposition"/>
  </f>
  <f  xml:id="fl-wr"  name="class">
    <symbol  value="pronoun"/>
  </f>
  <f  xml:id="fl-wv"  name="class">
    <symbol  value="verb"/>
  </f>
<!-- ... -->
</fLib>
```

Alternatively, a stand-off technique may be used, as in the following example, where a <linkGrp> element is used to link selected characters in the text *Caesar seized control* with their phonological representations.

```
<s>
  <w  xml:id="S1W1">
    <c  xml:id="S1W1C1">C</c>ae<c  xml:id="S1W1C2">s</c>ar</w>
  <w  xml:id="S1W2">
    <c  xml:id="S1W2C1">s</c>ei<c  xml:id="S1W2C2">z</c>e<c  xml:id="S1W2C3">d</c>
  </w>
  <w  xml:id="S1W3">con<c  xml:id="S1W3C1">t</c>rol</w>.

</s>
<fvLib  xml:id="FSL1"
  n="phonological segment definitions">
<!-- as in previous example -->
</fvLib>
<linkGrp  type="phonology">
<!-- ... -->
  <link  target="#S.DF #S1W3C1"/>
  <link  target="#Z.DF #S1W2C3"/>
  <link  target="#S.DF #S1W2C1"/>
  <link  target="#Z.DF #S1W2C2"/>
<!-- ... -->
</linkGrp>
```

As this example shows, a stand-off solution requires that every component to be linked to must be addressable in some way, by means of an XPointer. To handle the POS tagging example above, for example, each annotated element might be given an identifier of some sort, as follows:

```
<s  xml:id="mds09"  n="00741">
  <w  xml:id="mds0901">The</w>
  <w  xml:id="mds0902">closest</w>
  <w  xml:id="mds0903">he</w>
  <w  xml:id="mds0904">came</w>
  <w  xml:id="mds0905">to</w>
  <w  xml:id="mds0906">exercise</w>
<!-- ... -->
</s>
```

It would then be possible to link each word to its intended annotation in the feature library quoted above, as follows:

```
<linkGrp  type="POS-codes">
<!-- ... -->
  <link  target="#mds0901 #at0"/>
  <link  target="#mds0902 #ajs"/>
  <link  target="#mds0903 #pnp"/>
  <link  target="#mds0904 #vvd"/>
  <link  target="#mds0905 #prp"/>
  <link  target="#mds0906 #nn1"/>
  <link  target="#mds0907 #vbd"/>
  <link  target="#mds0908 #to0"/>
  <link  target="#mds0909 #vvi"/>
  <link  target="#mds0910 #crd"/>
<!-- ... -->
</linkGrp>
```

## 19.11 Feature System Declaration

The Feature System Declaration (FSD) is intended for use in conjunction with a TEI-conforming text that makes use of <fs> (that is, feature structure) elements. The FSD serves three purposes:

- the encoder can list all of the feature names and feature values and give a prose description as to what each represents.
- the encoder can define what it means to be a well-formed feature structure, and define constraints which may be used to determine whether a particular feature structure is *valid* relative to a given theory stated in typed feature logic. These may involve constraints on the range of a feature value, constraints on what features are valid within certain types of feature structures, or constraints that prevent the co-occurrence of certain feature-value pairs.
- the encoder can define the intended interpretation of underspecified feature structures. This involves defining default values (whether literal or computed) for missing features.

The scheme described in this chapter may be used to document any feature structure system, but is primarily intended for use with the feature structure representation defined by the ISO 24610-1:2006 standard, which corresponds with the recommendations presented in these Guidelines, *19. Feature Structures*. This chapter relies upon, but does not reproduce, formal definitions and descriptions presented more thoroughly in the ISO standard, which should be consulted in case of ambiguity or uncertainty.

The FSD serves an important function in documenting precisely what the encoder intended by the system of feature structure markup used in an XML-encoded text. The FSD is also an important resource which standardizes the rules of inference used by software to validate the feature structure markup in a text, and to infer the full interpretation of underspecified feature structures.

The reader should be aware the terminology used in this document does not always closely follow conventional practice in formal logic, and may also diverge from practice in some linguistic applications of typed feature structures. In particular, the term 'interpretation' when applied to a feature structure is not an interpretation in the model-theoretic sense, but is instead a minimally informative (or equivalently, most general) extension of that feature structure that is consistent with a set of constraints declared by an FSD. In linguistic application, such a system of constraints is the principal means by which the grammar of some natural language is expressed. There is a great deal of disagreement as to what, if any, model-theoretic interpretation feature structures have in such applications, but the status of this formal kind of interpretation is not germane to the present document. Similarly, the term 'valid' is used here as elsewhere in these Guidelines to identify the syntactic state of well-formedness in the sense defined by the logic of typed feature structures itself, as distinct from and in addition to the 'well-formedness' that pertains at the level of this encoding standard. No appeal to any notion from formal semantics should be inferred.

We begin by describing how an encoded text is associated with one or more feature system declarations. The second, third, and fourth sections describe the overall structure of a feature system declaration and give details of how to encode its components. The final section offers a full example; fuller discussion of the reasoning behind FSDs and another complete example are provided in **?**.

### 19.11.1 Linking a TEI Text to Feature System Declarations

In order for application software to use feature system declarations to aid in the automatic interpretation of encoded texts, or even for human readers to find the appropriate declarations which document the feature system used in markup, there must be a formal link from the encoded texts to the declarations. However, the schema which declares the syntax of the Feature System itself should be kept distinct from the feature structure schema, which is an application of that system.

A document containing typed feature structures may simply include a feature system declaration documenting those feature structures. A more usual scenario, however, is that the same feature system declaration (or parts of it) will be shared by many documents. In either case, an <fsDecl> element for each distinct type of feature structure used must be provided and associated with the type, which is the value used within each feature structure for its *type* attribute.

When the module defined in this chapter is included in an XML schema, the following elements become available via the `model.fsdDeclPart` class:

`<fsdDecl>` (feature system declaration) provides a feature system declaration comprising one or more feature structure declarations or feature structure declaration links.

`model.fsdDeclPart` groups elements which can occur as direct children of <fsdDecl>.

| | |
|---|---|
| fLib | (feature library) assembles a library of <f> (feature) elements. |
| fsDecl | (feature structure declaration) declares one type of feature structure. |
| fsdLink | (feature structure declaration link) associates the name of a typed feature structure with a feature structure declaration for it. |
| fvLib | (feature-value library) assembles a library of reusable feature value elements (including complete feature structures). |

The <fsdDecl> element serves as a wrapper for declaring feature systems and may be supplied either within the header of a standard TEI document, or as a standalone document in its own right. It contains one or more <fsdLink> or <fsDecl> elements and may hold several <fLib> or <fvLib> as well.

For example, suppose that a document `doc.xml` contains feature structures of two types: gpsg and lex. We might simply embed an <fsDecl> element for each within the header attached to the document as follows:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
<!-- example -->
    </fileDesc>
    <encodingDesc>
<!-- ... -->
      <fsdDecl>
        <fsDecl  type="gpsg">
<!-- information about this type -->
        </fsDecl>
        <fsDecl  type="lex">
<!-- information about this type -->
        </fsDecl>
      </fsdDecl>
<!-- ... -->
    </encodingDesc>
  </teiHeader>
  <text>
    <body>
<!-- ... -->
      <fs  type="lex">
<!-- an instance of the typed feature structure "lex" -->
      </fs>
```

```
<!-- ... -->
   </body>
  </text>
</TEI>
```

In this case there is an implicit link between the <fs> element and the corresponding <fsDecl> element because they share the same value for their *type* attribute and appear within the same document. This is a short cut for the more general case which requires a more explicit link provided by means of the <fsdLink> element, as demonstrated below.

Now suppose that we wish to create a second document which includes feature structures of the same type. Rather than duplicate the corresponding declarations, we will need to provide a means of pointing to them from this second document. The easiest[119] way of accomplishing this is to add an XML identifier to each <fsDecl> element in `example.xml`:

```
<fsdDecl>
  <fsDecl  type="gpsg"  xml:id="GPSG">
<!-- information about this type -->
  </fsDecl>
  <fsDecl  type="lex"  xml:id="LEX">
<!-- information about this type -->
  </fsDecl>
</fsdDecl>
```

(Although in this case the XML identifier is simply an uppercase version of the type name, there is no necessary connection between the two names. The only requirement is that the XML identifier conform to the standards required for identifiers, and that it be unique within the document containing it.)

In the <fsdDecl> for the second document, we can now include pointers to the <fsDecl> elements in the first:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
<!-- doc2 -->
    </fileDesc>
    <encodingDesc>
<!-- ... -->
      <fsdDecl>
        <fsdLink  type="gpsg"
          target="example.xml#GPSG"/>
        <fsdLink  type="lexx"
          target="example.xml#LEX"/>
      </fsdDecl>
<!-- ... -->
    </encodingDesc>
  </teiHeader>
  <text>
    <body>
<!-- ... -->
      <fs  type="lexx">
<!-- an instance of the typed feature structure "lex" -->
      </fs>
<!-- ... -->
    </body>
```

[119] Ways of pointing to components of a TEI document without using an XML identifier are discussed in *17.2.1. Pointing Elsewhere*

```
  </text>
</TEI>
```

Note that in `doc2.xml` there is no requirement for the local name for a given type of feature structures to be the same as that used by `example.xml`. We assume in this encoding that the type called lexx in `doc2.xml` is declared as having identical constraints and other properties to those declared for the type called lex in `example.xml`.

An <fsdDecl> may be given, as above, within the encoding description of the <teiHeader> element of a TEI document containing typed feature structures. Alternatively, it may appear independently of any feature structures, as a document in its own right with its own <teiHeader>. These options are both possible because the element is a member of both the `model.encodingDescPart` class and the `model.resource` class.

The current recommendations provide no way of enforcing uniqueness of the *type* values among <fsdDecl> elements, nor of requiring that every *type* value specified on an <fs> element be also declared on an <fsdDecl> element. Encoders requiring such constraints (which might have some obvious utility in assisting the consistency and accuracy of tagging) are recommended to develop tools to enforce them, using such mechanisms as Schematron assertions.

### 19.11.2 The Overall Structure of a Feature System Declaration

A feature system declaration contains one or more feature structure declarations, each of which has up to three parts: an optional description (which gives a prose comment on what that type of feature structure encodes), an obligatory set of feature declarations (which specify range constraints and default values for the features in that type of structure), and optional feature structure constraints (which specify co-occurrence restrictions on feature values).

`<fsDescr>` (feature system description (in FSD)) describes in prose what is represented by the type of feature structure declared in the enclosing fsDecl.

`<fDecl>` (feature declaration) declares a single feature, specifying its name, organization, range of allowed values, and optionally its default value.

`<fsConstraints>` (feature-structure constraints) specifies constraints on the content of valid feature structures.

Feature declarations and feature structure constraints are described in the next two sections. Note that the specification of similar <fsDecl> elements can be simplified by devising an inheritance hierarchy for the feature structure types. Each <fsDecl> element may name one or more 'basetypes' from which it inherits feature declarations and constraints (these are often called 'supertypes'). For instance, suppose that <fsDecl type="Basic"> contains <fDecl name="One"> and <fDecl name="Two">, and that <fsDecl type="Derived" baseTypes="Basic"> contains just <fDecl name="Three">. Then any instance of <fs type="Derived"> must include all three features. This is because <fsDecl type="Derived"> inherits the two feature declarations from <fsDecl type="Basic"> when it specifies a base type of Basic.

The following sample shows the overall structure of a complete feature structure declaration:

```
<fsDecl type="SomeName">
  <fsDescr>Describes what this type of fs represents</fsDescr>
  <fDecl name="featureOne">
<!-- The declaration for featureOne -->
  </fDecl>
  <fDecl name="featureTwo">
<!-- The declaration for featureTwo -->
  </fDecl>
  <fsConstraints>
<!-- The feature structure constraints go here -->
  </fsConstraints>
</fsDecl>
```

The attribute *baseTypes* gives the name of one or more types from which this type inherits feature specifications and constraints; if this type includes a feature specification with the same name as one inherited from any of the

types specified by this attribute, or if more than one specification of the same name is inherited, then the possible values of that feature is determined by unification. Similarly, the set of constraints applicable is derived by conjoining those specified explicitly within this element with those implied by the *baseTypes* attribute. When no base type is specified, no feature specification or constraint is inherited.

Although the present standard does provide for default feature values, feature inheritance is defined to be monotonic.

The process of combining constraints may result in a contradiction, for example if two specifications for the same feature specify disjoint ranges of values, and at least one such specification is mandatory. In such a case, there is no valid feature structure of the type being defined.

Every type specified by *baseTypes* must be a single word which is a legal XML name; for example, they cannot include whitespace or begin with digits. Multiple base types are separated with spaces, e.g. <fsDecl type="Sub" baseTypes="Super1 Super2">.

### 19.11.3 Feature Declarations

Each feature is declared in an <fDecl> element whose *name* attribute identifies the feature being declared; this matches the *name* attribute of the <f> elements it declares. An <fDecl> has three parts: an optional prose description (which should explain what the feature and its values represent), an obligatory range specification (which declares what values the feature is allowed to have), and an optional default specification (which declares what default value should be supplied when the named feature does not appear in an <fs>). If, in a feature structure, a feature:

- is not optional (i.e., is obligatory),
- has no value provided, or the value <default> is provided (see ISO 24610-1, Subclause 5.10, Default Values, and
- either has no default specified, or has conditional defaults, none of the conditions on which is met,

then the value of this feature in the feature structure's most general valid extension is the most general value provided in its <vRange>, in the case of a unit organization, or the singleton set, bag, or list containing that element, in the case of a complex organization. If the feature:

- is optional,
- has no value provided, or the value <default> is provided, and
- either has a default specified, or has conditional defaults, one of the conditions on which is met,

then this feature does have a value in the feature structure's most general valid extension when it exists, namely the default value that pertains.

It is possible that a feature structure will not have a valid extension because the default value that pertains to a feature is not consistent with that feature's declared range. Additional tools are required for the enforcement of such criteria.

The following elements are used in feature system declarations:

`<fDecl>` (feature declaration) declares a single feature, specifying its name, organization, range of allowed values, and optionally its default value.

`@name` a single word which follows the rules defining a legal XML name (see `https://www.w3.org/TR/REC-xml/#dt-name`), indicating the name of the feature being declared; matches the *name* attribute of <f> elements in the text.

`@optional` indicates whether or not the value of this feature may be present.

`<fDescr>` (feature description (in FSD)) describes in prose what is represented by the feature being declared and its values.

`<vRange>` (value range) defines the range of allowed values for a feature, in the form of an <fs>, <vAlt>, or primitive value; for the value of an <f> to be valid, it must be subsumed by the specified range; if the <f> contains multiple values (as sanctioned by the *org* attribute), then each value must be subsumed by the <vRange>.

`<vDefault>` (value default) declares the default value to be supplied when a feature structure does not contain an instance of <f> for this name; if unconditional, it is specified as one (or, depending on the value of the *org* attribute of the enclosing <fDecl>) more <fs> elements or primitive values; if conditional,

it is specified as one or more <if> elements; if no default is specified, or no condition matches, the value none is assumed.

`<if>` defines a conditional default value for a feature; the condition is specified as a feature structure, and is met if it subsumes the feature structure in the text for which a default value is sought.

`<then>` separates the condition from the default in an <if>, or the antecedent and the consequent in a <cond> element.

The logic for validating feature values and for matching the conditions for supplying default values is based on the operation of *subsumption*. Subsumption is a standard operation in feature-structure-based formalisms. Informally, a feature structure FS subsumes all feature structures that are at least as informative as itself; that is, all feature structures that specify all of the feature values that FS does with values that are subsumed by the values that FS has, and that have all of the re-entrancies (see *19.6. Re-entrant Feature Structures*) that FS does. (**?**; see also **?** and **?**) A more formal definition is provided in ISO 24610-1:2006 .

Following the spirit of the informal definition above, we can extend subsumption in a straightforward way to cover alternation, negation, special primitive values, and the use of attributes in the markup. For instance, a <vAlt> containing the value v subsumes v. The negation of a value v (represented by means of the <vNot> element discussed in section *19.8.2. Negation*) subsumes any value that is not v; for example `<vNot><numeric value='0'/></vNot>` subsumes any numeric value other than zero. The value <fs type="X"/> subsumes any feature structure of type X, even if it is not valid.

As an example of feature declarations, consider the following extract from Gazdar et al.'s *Generalized Phrase Structure Grammar*. In the appendix to their book, they propose a feature system for English of which this is just a sampling:

```
feature    value range
INV        {+, -}
CONJ       {and, both, but, either, neither, nor, or, NIL}
COMP       {for, that, whether, if, NIL}
AGR        CAT
PFORM      {to, by, for, ...}
```

```
Feature specification defaults
FSD 1:  [-INV]
FSD 2:  ~[CONJ]
FSD 9:  [INF, +SUBJ] --> [COMP for]
```

The INV feature, which encodes whether or not a sentence is inverted, allows only the values plus (+) and minus (-). If the feature is not specified, then the default rule (FSD 1 above) says that a value of minus is always assumed. The feature declaration for this feature would be encoded as follows:

```
<fDecl name="INV">
  <fDescr>inverted sentence</fDescr>
  <vRange>
    <vAlt>
      <binary value="true"/>
      <binary value="false"/>
    </vAlt>
  </vRange>
  <vDefault>
    <binary value="false"/>
  </vDefault>
</fDecl>
```

The value range is specified as an alternation (more precisely, an exclusive disjunction), which can be represented by the <binary> feature value. That is, the value must be either true or false, but cannot be both or neither.

The CONJ feature indicates the surface form of the conjunction used in a construction. The ~in the default rule (see FSD 2 above) represents negation. This means that by default the feature is not applicable, in other words, no conjunction is taking place. Note that CONJ not being present is distinct from CONJ being present but having the NIL value allowed in the value range. In their analysis, NIL means that the phenomenon of conjunction is taking place but there is no explicit conjunction in the surface form of the sentence. The feature declaration for this feature would be encoded as follows:

```
<fDecl name="CONJ">
  <fDescr>surface form of the conjunction</fDescr>
  <vRange>
    <vAlt>
      <symbol value="and"/>
      <symbol value="both"/>
      <symbol value="but"/>
      <symbol value="either"/>
      <symbol value="neither"/>
      <symbol value="nor"/>
      <symbol value="or"/>
      <symbol value="NIL"/>
    </vAlt>
  </vRange>
  <vDefault>
    <binary value="false"/>
  </vDefault>
</fDecl>
```

Note that the <vDefault> is not strictly necessary in this case, since the binary value of false only serves to convey the information that the feature has no other legitimate value.

The COMP feature indicates the surface form of the complementizer used in a construction. In value range, it is analogous to CONJ. However, its default rule (see FSD 9 above) is conditional. It says that if the verb form is infinitival (the VFORM feature is not mentioned in the rule since it is the only feature that can take INF as a value), and the construction has a subject, then a *for* complement must be used. For instance, to make John the subject of the infinitive in *It is necessary to go,* a *for* complement must be used; that is, *It is necessary for John to go.* The feature declaration for this feature would be encoded as follows:

```
<fDecl name="COMP">
  <fDescr>surface form of the complementizer</fDescr>
  <vRange>
    <vAlt>
      <symbol value="for"/>
      <symbol value="that"/>
      <symbol value="whether"/>
      <symbol value="if"/>
      <symbol value="NIL"/>
    </vAlt>
  </vRange>
  <vDefault>
    <if>
      <fs>
        <f name="VFORM">
          <symbol value="INF"/>
        </f>
        <f name="SUBJ">
```

```
        <binary value="true"/>
      </f>
    </fs>
    <then/>
    <symbol value="for"/>
  </if>
 </vDefault>
</fDecl>
```

The AGR feature stores the features relevant to subject-verb agreement. Gazdar et al. specify the range of this feature as CAT. This means that the value is a *category*, which is their term for a feature structure. This is actually too weak a statement. Not just any feature structure is allowable here; it must be a feature structure for agreement (which is defined in the complete example at the end of the chapter to contain the features of person and number). The following feature declaration encodes this constraint on the value range:

```
<fDecl name="AGR">
 <fDescr>agreement for person and number</fDescr>
 <vRange>
  <fs type="Agreement"/>
 </vRange>
</fDecl>
```

That is, the value must be a feature structure of type Agreement. The complete example at the end of this chapter includes the <fsDecl type="Agreement"> which includes <fDecl name="PERS"> and <fDecl name="NUM">.

The PFORM feature indicates the surface form of the preposition used in a construction. Since PFORM is specified above as an open set, <string> is used in the range specification below rather than <symbol>.

```
<fDecl name="PFORM">
 <fDescr>word form of a preposition</fDescr>
 <vRange>
  <vNot>
   <string/>
  </vNot>
 </vRange>
</fDecl>
```

This example makes use of a negated value: `<vNot><string/></vNot>` subsumes any string that is not the empty string.

For the reduced feature structure that concludes section *19.3. Other Atomic Feature Values* above and identifies the value of some part of speech to be a common noun, it is possible to align the concept of *part of speech* with its definition and persistent identifier using the *targetDatcat* attribute, which connects the modeled XML object with the appropriate locus in a reference taxonomy, as shown below:

```
<fDecl name="POS"
 targetDatcat="http://hdl.handle.net/11459/CCR_C-396_5a972b93-2294-ab5c-a541-7c344c5f26c3">
 <fDescr>part of speech (morphosyntactic category)</fDescr>
 <vRange>
  <vAlt>
   <symbol value="NN"
    datcat="http://hdl.handle.net/11459/CCR_C-1256_7ec6083c-23d4-224d-6f94-eecbe6861545"/>
   <symbol value="NP"
    datcat="http://hdl.handle.net/11459/CCR_C-1371_fbebd9ec-a7f4-9a36-d6e9-88ee16b944ae"/>
```

```
<!-- ... -->
    </vAlt>
  </vRange>
</fDecl>
```

The above example declares the feature 'POS' as instantiating the corresponding concept defined in a reference taxonomy or ontology, and defines the range of values of the feature at hand by listing the appropriate alternatives, together with their external persistent identifiers.

Note that the class `model.featureVal` includes all possible single feature values, including feature structures, alternations (<vAlt>) and complex collections (<vColl>).

### 19.11.4 Feature Structure Constraints

Ensuring the validity of feature structures may require much more than simply specifying the range of allowed values for each feature. There may be constraints on the co-occurrence of one feature value with the value of another feature in the same feature structure or in an embedded feature structure.

Such constraints on valid feature structures are expressed as a series of conditional and biconditional tests in the <fsConstraints> part of an <fsDecl>. A particular feature structure is valid only if it meets all the constraints. The <cond> element encodes the conventional if-then conditional of boolean logic which succeeds when both the antecedent and consequent are true, or whenever the antecedent is false. The <bicond> element encodes the biconditional (if and only if) operation of boolean logic. It succeeds only when the corresponding if-then conditionals in both directions are true. In feature structure constraints the antecedent and consequent are expressed as feature structures; they are considered true if they *subsume* (see section *19.11.3. Feature Declarations*) the feature structure in question, but in the case of consequents, this truth is asserted rather than simply tested. That is to say, a conditional is enforced by determining that the antecedent does not (and will never) subsume the given feature structure, or by determining that the antecedent does subsume the given feature structure, and then unifying the consequent with it (the result of which, if successful, will be subsumed by the consequent). In practice, the enforcement of such constraints can result in periods in which the truth of a constraint with respect to a given feature structure is simply not known; in this case, the constraint must be persistently monitored as the feature structure becomes more informative until either its truth value is determined or computation fails for some other reason.

The following elements make up the <fsConstraints> part of an FSD:

`<fsConstraints>` (feature-structure constraints) specifies constraints on the content of valid feature structures.

`<cond>` (conditional feature-structure constraint) defines a conditional feature-structure constraint; the consequent and the antecedent are specified as feature structures or feature-structure collections; the constraint is satisfied if both the antecedent and the consequent subsume a given feature structure, or if the antecedent does not.

`<bicond>` (bi-conditional feature-structure constraint) defines a biconditional feature-structure constraint; both consequent and antecedent are specified as feature structures or groups of feature structures; the constraint is satisfied if both subsume a given feature structure, or if both do not.

`<then>` separates the condition from the default in an <if>, or the antecedent and the consequent in a <cond> element.

`<iff>` (if and only if) separates the condition from the consequence in a <bicond> element.

For an example of feature structure constraints, consider the following 'feature co-occurrence restrictions' extracted from the feature system for English proposed by Gazdar, et al. (1985:246–247):

```
[FCR 1: [+INV] → [+AUX, FIN]
```

```
FCR 7: [BAR 0] ≡ [N] & [V] & [SUBCAT]
```

```
FCR 8: [BAR 1] → ~ [SUBCAT]]
```

The first constraint says that if a construction is inverted, it must also have an auxiliary and a finite verb form. That is,

```
<cond>
  <fs>
    <f  name="INV">
      <binary  value="true"/>
    </f>
  </fs>
  <then/>
  <fs>
    <f  name="AUX">
      <binary  value="true"/>
    </f>
    <f  name="VFORM">
      <symbol  value="FIN"/>
    </f>
  </fs>
</cond>
```

The second constraint says that if a construction has a BAR value of zero (i.e., it is a sentence), then it must have a value for the features N, V, and SUBCAT. By the same token, because it is a biconditional, if it has values for N, V, and SUBCAT, it must have BAR='0'. That is,

```
<bicond>
  <fs>
    <f  name="BAR">
      <symbol  value="0"/>
    </f>
  </fs>
  <iff/>
  <fs>
    <f  name="N">
      <binary  value="true"/>
    </f>
    <f  name="V">
      <binary  value="true"/>
    </f>
    <f  name="SUBCAT">
      <binary  value="true"/>
    </f>
  </fs>
</bicond>
```

The final constraint says that if a construction has a BAR value of 1 (i.e., it is a phrase), then the SUBCAT feature should be absent (~). This is not biconditional, since there are other instances under which the SUBCAT feature is inappropriate. That is,

```
<cond>
  <fs>
    <f  name="BAR">
      <symbol  value="1"/>
    </f>
  </fs>
  <then/>
  <fs>
    <f  name="SUBCAT">
      <binary  value="false"/>
    </f>
  </fs>
</cond>
```

Note that <cond> and <bicond> use the empty tags <then> and <iff>, respectively, to separate the antecedent and consequent. These are primarily for the sake of enhancing human readability.

### 19.11.5 A Complete Example

To summarize this chapter, the complete FSD for the example that has run through the chapter is reproduced below:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>A sample FSD based on an extract from Gazdar
                et al.'s GPSG feature system for English</title>
        <respStmt>
          <resp>encoded by</resp>
          <name>Gary F. Simons</name>
        </respStmt>
      </titleStmt>
      <publicationStmt>
        <p>This sample was first encoded by Gary F. Simons (Summer
                Institute of Linguistics, Dallas, TX) on January 28, 1991.
                Revised April 8, 1993 to match the specification of FSDs
                in version P2 of the TEI Guidelines. Revised again December 2004 to
                be consistent with the feature structure representation standard
                jointly developed with ISO TC37/SC4.
        </p>
      </publicationStmt>
      <sourceDesc>
        <p>This sample FSD does not describe a complete feature
                system. It is based on extracts from the feature system
                for English presented in the appendix (pages 245–247) of
                Generalized Phrase Structure Grammar, by Gazdar, Klein,
                Pullum, and Sag (Harvard University Press, 1985).</p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <fsdDecl>
    <fsDecl  type="GPSG">
      <fsDescr>Encodes a feature structure for the GPSG analysis
            of English (after Gazdar, Klein, Pullum, and Sag)</fsDescr>
      <fDecl  name="INV">
        <fDescr>inverted sentence</fDescr>
        <vRange>
```

```
      <vAlt>
        <binary value="true"/>
        <binary value="false"/>
      </vAlt>
    </vRange>
    <vDefault>
      <binary value="false"/>
    </vDefault>
  </fDecl>
  <fDecl name="CONJ">
    <fDescr>surface form of the conjunction</fDescr>
    <vRange>
      <vAlt>
        <symbol value="and"/>
        <symbol value="both"/>
        <symbol value="but"/>
        <symbol value="either"/>
        <symbol value="neither"/>
        <symbol value="nor"/>
        <symbol value="or"/>
        <symbol value="NIL"/>
      </vAlt>
    </vRange>
    <vDefault>
      <binary value="false"/>
    </vDefault>
  </fDecl>
  <fDecl name="COMP">
    <fDescr>surface form of the complementizer</fDescr>
    <vRange>
      <vAlt>
        <symbol value="for"/>
        <symbol value="that"/>
        <symbol value="whether"/>
        <symbol value="if"/>
        <symbol value="NIL"/>
      </vAlt>
    </vRange>
    <vDefault>
      <if>
        <fs>
          <f name="VFORM">
            <symbol value="INF"/>
          </f>
          <f name="SUBJ">
            <binary value="true"/>
          </f>
        </fs>
        <then/>
        <symbol value="for"/>
      </if>
    </vDefault>
  </fDecl>
  <fDecl name="AGR">
    <fDescr>agreement for person and number</fDescr>
    <vRange>
      <fs type="Agreement"/>
```

```
    </vRange>
  </fDecl>
  <fDecl name="PFORM">
    <fDescr>word form of a preposition</fDescr>
    <vRange>
      <vNot>
        <string/>
      </vNot>
    </vRange>
  </fDecl>
  <fsConstraints>
    <cond>
      <fs>
        <f name="INV">
          <binary value="true"/>
        </f>
      </fs>
      <then/>
      <fs>
        <f name="AUX">
          <binary value="true"/>
        </f>
        <f name="VFORM">
          <symbol value="FIN"/>
        </f>
      </fs>
    </cond>
    <bicond>
      <fs>
        <f name="BAR">
          <symbol value="0"/>
        </f>
      </fs>
      <iff/>
      <fs>
        <f name="N">
          <binary value="true"/>
        </f>
        <f name="V">
          <binary value="true"/>
        </f>
        <f name="SUBCAT">
          <binary value="true"/>
        </f>
      </fs>
    </bicond>
    <cond>
      <fs>
        <f name="BAR">
          <symbol value="1"/>
        </f>
      </fs>
      <then/>
      <fs>
        <f name="SUBCAT">
          <binary value="false"/>
        </f>
```

```
        </fs>
      </cond>
    </fsConstraints>
  </fsDecl>
  <fsDecl type="Agreement">
    <fsDescr>This type of feature structure encodes the features
          for subject-verb agreement in English</fsDescr>
    <fDecl name="PERS">
      <fDescr>person (first, second, or third)</fDescr>
      <vRange>
        <vAlt>
          <symbol value="1"/>
          <symbol value="2"/>
          <symbol value="3"/>
        </vAlt>
      </vRange>
    </fDecl>
    <fDecl name="NUM">
      <fDescr>number (singular or plural)</fDescr>
      <vRange>
        <vAlt>
          <symbol value="sg"/>
          <symbol value="pl"/>
        </vAlt>
      </vRange>
    </fDecl>
  </fsDecl>
 </fsdDecl>
</TEI>
```

## 19.12 Formal Definition and Implementation

This elements discussed in this chapter constitute a module of the TEI scheme which is formally defined as follows:

**Module iso-fs: Feature structures**

- *Elements defined*: bicond binary cond default f fDecl fDescr fLib fs fsConstraints fsDecl fsDescr fsdDecl fsdLink fvLib if iff numeric string symbol then vAlt vColl vDefault vLabel vMerge vNot vRange
- *Classes defined*: model.fsdDeclPart

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 20 Graphs, Networks, and Trees

Graphical representations are widely used for displaying relations among informational units because they help readers to visualize those relations and hence to understand them better. Two general types of graphical representations may be distinguished.

- *Graphs*, in the strictly mathematical sense, consist of points, often called *nodes* or *vertices*, and connections among them, called *arcs*, or under certain conditions, *edges*. Among the various types of graphs are *networks* and *trees*. Graphs generally and networks in particular are dealt with directly below. Trees are dealt with separately in sections *20.2. Trees* and *20.3. Another Tree Notation*.[120]
- *Charts*, which typically plot data in two or more dimensions, including plots with orthogonal or radial axes, bar charts, pie charts, and the like. These can be described using the elements defined in the module for figures and graphics; see chapter *15. Tables, Formulæ, Graphics, and Notated Music*.

Among the types of qualitative relations often represented by graphs are organizational hierarchies, flow charts, genealogies, semantic networks, transition networks, grammatical relations, tournament schedules, seating plans, and directions to people's houses. In developing recommendations for the encoding of graphs of various types, we have relied on their formal mathematical definitions and on the most common conventions for representing them visually. However, it must be emphasized that these recommendations do not provide for the full range of possible graphical representations, and deal only partially with questions of design, layout, and placement.

## 20.1 Graphs and Digraphs

Broadly speaking, graphs can be divided into two types: *undirected* and *directed*. An undirected graph is a set of *nodes* (or *vertices*) together with a set of pairs of those vertices, called *arcs* or *edges*. Each node in an arc of an undirected graph is said to be *incident* with that arc, and the two vertices (nodes) which make up an arc are said to be *adjacent*. An directed graph is like an undirected graph except that the arcs are *ordered pairs* of nodes. In the case of directed graphs, the term *edge* is not used; moreover, each arc in a directed graph is said to be *adjacent from* the node from which the arc emanates, and *adjacent to* the node to which the arc is directed. We use the element <graph> to encode graphs as a whole, <node> to encode nodes or vertices, and <arc> to encode arcs or edges; arcs can also be encoded by attributes on the <node> element. These elements have the following descriptions and attributes:

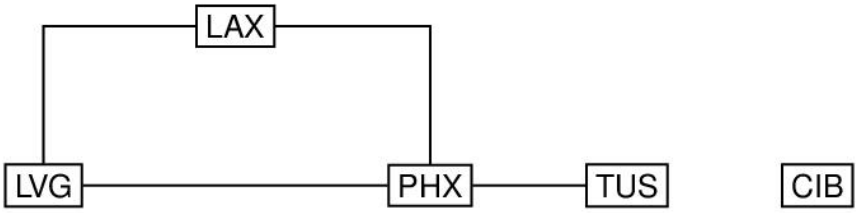
`<graph>` (graph) encodes a graph, which is a collection of nodes, and arcs which connect the nodes.

`<node>` (node) encodes a node, a possibly labeled point in a graph.

`<arc>` (arc) encodes an arc, the connection from one node to another in a graph.

Before proceeding, some additional terminology may be helpful. We define a *path* in a graph as a sequence of nodes n1, ..., nk such that there is an arc from each ni to ni+1 in the sequence. A *cyclic path*, or *cycle* is a path leading from a particular node back to itself. A graph that contains at least one cycle is said to be *cyclic*; otherwise it is *acyclic*. We say, finally, that a graph is *connected* if there is a path from some node to every other node in the graph; any graph that is not connected is said to be *disconnected*.

Here is an example of an undirected, cyclic disconnected graph, in which the nodes are annotated with three-letter codes for airports, and the arcs connecting the nodes are represented by horizontal and vertical lines, with 90 degree bends used simply to avoid having to draw diagonal lines.




Next is a markup of the graph, using <arc> elements to encode the arcs.

```
<graph  type="undirected"  xml:id="CUG1"
  order="5"  size="4">
```

[120] The treatment here is largely based on the characterizations of graph types in ?

```
  <label>Airline Connections in Southwestern USA</label>
  <node  xml:id="LAX"  degree="2">
    <label>LAX</label>
  </node>
  <node  xml:id="LVG"  degree="2">
    <label>LVG</label>
  </node>
  <node  xml:id="PHX"  degree="3">
    <label>PHX</label>
  </node>
  <node  xml:id="TUS"  degree="1">
    <label>TUS</label>
  </node>
  <node  xml:id="CIB"  degree="0">
    <label>CIB</label>
  </node>
  <arc  from="#LAX"  to="#LVG"/>
  <arc  from="#LAX"  to="#PHX"/>
  <arc  from="#LVG"  to="#PHX"/>
  <arc  from="#PHX"  to="#TUS"/>
</graph>
```

The first child element of <graph> may be a <label> to record a label for the graph; similarly, the <label> child of each <node> element records the labels of that node. The *order* and *size* attributes on the <graph> element record the number of nodes and number of arcs in the graph respectively; these values are optional (since they can be computed from the rest of the graph), but if they are supplied, they must be consistent with the rest of the encoding. They can thus be used to help check that the graph has been encoded and transmitted correctly. The *degree* attribute on the <node> elements record the number of arcs that are incident with that node. It is optional (because redundant), but can be used to help in validity checking: if a value is given, it must be consistent with the rest of the information in the graph. Finally, the *from* and *to* attributes on the <arc> elements provide pointers to the nodes connected by those arcs. Since the graph is undirected, no directionality is implied by the use of the *from* and *to* attributes; the values of these attributes could be interchanged in each arc without changing the graph.

The *adj*, *adjFrom*, and *adjTo* attributes of the <node> element provide an alternative method of representing unlabeled arcs, their values being pointers to the nodes which are adjacent to or from that node. The *adj* attribute is to be used for undirected graphs, and the *adjFrom* and *adjTo* attributes for directed graphs. It is a semantic error for the directed adjacency attributes to be used in an undirected graph, and vice versa. Here is a markup of the preceding graph, using the *adj* attribute to represent the arcs.

```
<graph  type="undirected"  xml:id="CUG2"
  order="5"  size="4">
  <label>Airline Connections in Southwestern USA</label>
  <node  xml:id="LAX2"  degree="2"
    adj="#LVG2 #PHX2">
    <label>LAX2</label>
  </node>
  <node  xml:id="LVG2"  degree="2"
    adj="#LAX2 #PHX2">
    <label>LVG2</label>
  </node>
  <node  xml:id="PHX2"  degree="3"
    adj="#LAX2 #LVG2 #TUS2">
    <label>PHX2</label>
  </node>
  <node  xml:id="TUS2"  degree="1"  adj="#PHX2">
    <label>TUS2</label>
```

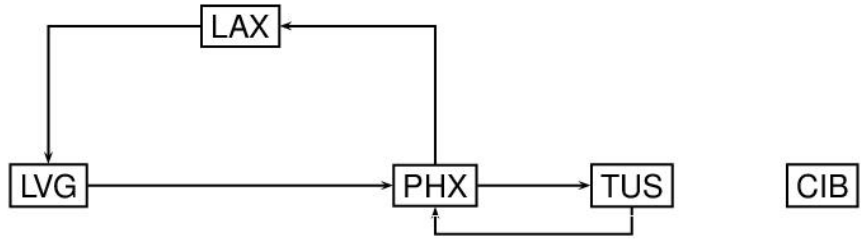
```
  </node>
  <node  xml:id="CIB2"  degree="0">
    <label>CIB2</label>
  </node>
</graph>
```

Note that each arc is represented twice in this encoding of the graph. For example, the existence of the arc from LAX to LVG can be inferred from each of the first two <node> elements in the graph. This redundancy, however, is not required: it suffices to describe an arc in any one of the three places it can be described (either adjacent node, or in a separate <arc> element). Here is a less redundant representation of the same graph.

```
<graph  type="undirected"  xml:id="CUG3"
  order="5"  size="4">
  <label>Airline Connections in Southwestern USA</label>
  <node  xml:id="LAX3"  degree="2"
    adj="#LVG3 #PHX3">
    <label>LAX3</label>
  </node>
  <node  xml:id="LVG3"  degree="2"  adj="#PHX3">
    <label>LVG3</label>
  </node>
  <node  xml:id="PHX3"  degree="3"  adj="#TUS3">
    <label>PHX3</label>
  </node>
  <node  xml:id="TUS3"  degree="1">
    <label>TUS3</label>
  </node>
  <node  xml:id="CIB3"  degree="0">
    <label>CIB3</label>
  </node>
</graph>
```

Although in many cases the <arc> element is redundant (since arcs can be described using the adjacency attributes of their adjacent nodes), it has nevertheless been included in this module, in order to allow the convenient specification of identifiers, display or rendition information, and labels for each arc (using the attributes *xml:id*, *rend*, and a child <label> element).

Next, let us modify the preceding graph by adding directionality to the arcs. Specifically, we now think of the arcs as specifying selected routes from one airport to another, as indicated by the direction of the arrowheads in the following diagram.




Here is an encoding of this graph, using the <arc> element to designate the arcs.

```
<graph  type="directed"  xml:id="RDG1"
  order="5"  size="5">
```

```
  <label>Selected Airline Routes in Southwestern USA</label>
  <node xml:id="LAX4" inDegree="1"
    outDegree="1">
    <label>LAX4</label>
  </node>
  <node xml:id="LVG4" inDegree="1"
    outDegree="1">
    <label>LVG4</label>
  </node>
  <node xml:id="PHX4" inDegree="2"
    outDegree="2">
    <label>PHX4</label>
  </node>
  <node xml:id="TUS4" inDegree="1"
    outDegree="1">
    <label>TUS4</label>
  </node>
  <node xml:id="CIB4" inDegree="0"
    outDegree="0">
    <label>CIB4</label>
  </node>
  <arc from="#LAX4" to="#LVG4"/>
  <arc from="#LVG4" to="#PHX4"/>
  <arc from="#PHX4" to="#LAX4"/>
  <arc from="#PHX4" to="#TUS4"/>
  <arc from="#TUS4" to="#PHX4"/>
</graph>
```

The attributes *inDegree* and *outDegree* indicate the number of nodes which are adjacent to and from the node concerned respectively.

Here is another encoding of the graph, using the *adjTo* and *adjFrom* attributes on nodes to designate the arcs.

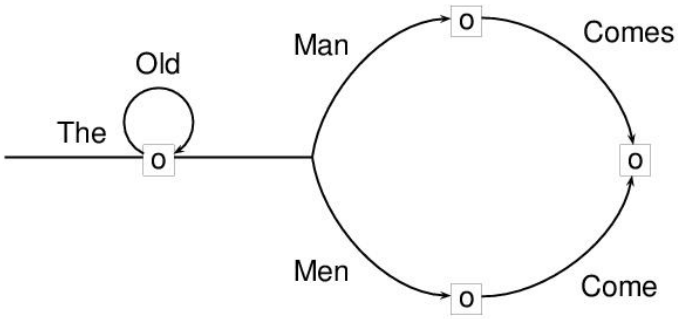
```
<graph type="directed" xml:id="RDG2"
  order="5" size="5">
  <label>Selected Airline Routes in Southwestern USA</label>
  <node xml:id="LAX5" inDegree="1"
    outDegree="1" adjTo="#LVG5" adjFrom="#PHX5">
    <label>LAX5</label>
  </node>
  <node xml:id="LVG5" inDegree="1"
    outDegree="1" adjFrom="#LAX5" adjTo="#PHX5">
    <label>LVG5</label>
  </node>
  <node xml:id="PHX5" inDegree="2"
    outDegree="2" adjTo="#LAX5 #TUS" adjFrom="#LVG5 #TUS5">
    <label>PHX5</label>
  </node>
  <node xml:id="TUS5" inDegree="1"
    outDegree="1" adjTo="#PHX5" adjFrom="#PHX5">
    <label>TUS5</label>
  </node>
  <node xml:id="CIB5" inDegree="0"
    outDegree="0">
    <label>CIB5</label>
  </node>
</graph>
```

If we wish to label the arcs, say with flight numbers, then <arc> elements must be used to hold the <label> elements, as in the following example.

```
<graph type="directed" xml:id="RDG3"
  order="5" size="5">
  <label>Selected Airline Routes in Southwestern USA</label>
  <node xml:id="LAX6">
    <label>LAX6</label>
  </node>
  <node xml:id="LVG6">
    <label>LVG6</label>
  </node>
  <node xml:id="PHX6">
    <label>PHX6</label>
  </node>
  <node xml:id="TUS6">
    <label>TUS6</label>
  </node>
  <node xml:id="CIB6">
    <label>CIB6</label>
  </node>
  <arc from="#LAX6" to="#LVG6">
    <label>SW117</label>
  </arc>
  <arc from="#LVG6" to="#PHX6">
    <label>SW711</label>
  </arc>
  <arc from="#PHX6" to="#LAX6">
    <label>AA218</label>
  </arc>
  <arc from="#PHX6" to="#TUS6">
    <label>AW229</label>
  </arc>
  <arc from="#TUS6" to="#PHX6">
    <label>AW225</label>
  </arc>
</graph>
```

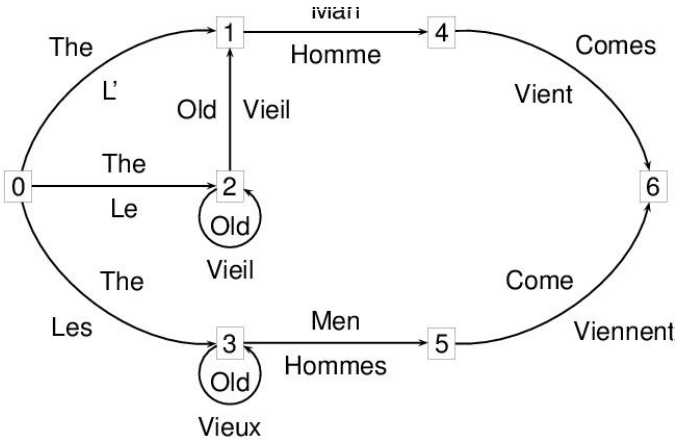
### 20.1.1 Transition Networks

For encoding transition networks and other kinds of directed graphs in which distinctions among types of nodes must be made, the *type* attribute is provided for <node> elements. In the following example, the *initial* and *final* nodes (or *states*) of the network are distinguished. It can be understood as accepting the set of strings obtained by traversing it from its initial node to its final node, and concatenating the labels.




```
<graph type="network-transition"
  xml:id="SS8" order="5" size="6">
  <label>(8)</label>
  <node xml:id="Q0" inDegree="0"
    outDegree="1" type="initial"/>
  <node xml:id="Q1" inDegree="2"
    outDegree="3"/>
  <node xml:id="Q2" inDegree="1"
    outDegree="1"/>
  <node xml:id="Q3" inDegree="1"
    outDegree="1"/>
  <node xml:id="Q4" inDegree="2"
    outDegree="0" type="final"/>
  <arc from="#Q0" to="#Q1">
    <label>THE</label>
  </arc>
  <arc from="#Q1" to="#Q1">
    <label>OLD</label>
  </arc>
  <arc from="#Q1" to="#Q2">
    <label>MAN</label>
  </arc>
  <arc from="#Q1" to="#Q3">
    <label>MEN</label>
  </arc>
  <arc from="#Q2" to="#Q4">
    <label>COMES</label>
  </arc>
  <arc from="#Q3" to="#Q4">
    <label>COME</label>
  </arc>
</graph>
```

A finite state transducer has two labels on each arc, and can be thought of as representing a mapping from one sequence of labels to the other. The following example represents a transducer for translating the English strings accepted by the network in the preceding example into French. The nodes have been annotated with numbers, for convenience.



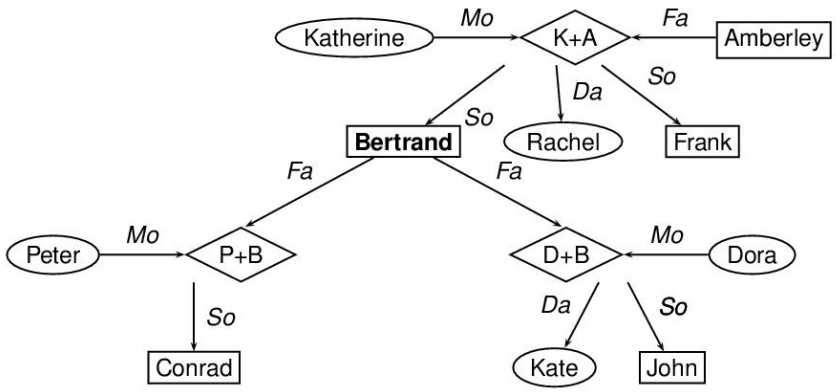

```
<graph type="transducer" order="7" size="10">
  <node xml:id="T0" inDegree="0"
    outDegree="3" type="initial">
    <label>0</label>
  </node>
  <node xml:id="T1" inDegree="2"
    outDegree="1">
    <label>1</label>
  </node>
  <node xml:id="T2" inDegree="2"
    outDegree="2">
    <label>2</label>
  </node>
  <node xml:id="T3" inDegree="2"
    outDegree="2">
    <label>3</label>
  </node>
  <node xml:id="T4" inDegree="1"
    outDegree="1">
    <label>4</label>
  </node>
  <node xml:id="T5" inDegree="1"
    outDegree="1">
    <label>5</label>
  </node>
  <node xml:id="T6" inDegree="2"
    outDegree="0" type="final">
    <label>6</label>
  </node>
  <arc from="#T0" to="#T1">
    <label>THE</label>
    <label>L'</label>
  </arc>
  <arc from="#T0" to="#T2">
    <label>THE</label>
    <label>LE</label>
  </arc>
  <arc from="#T0" to="#T3">
    <label>THE</label>
```

```
    <label>LES</label>
  </arc>
  <arc  from="#T1"  to="#T4">
    <label>MAN</label>
    <label>HOMME</label>
  </arc>
  <arc  from="#T2"  to="#T1">
    <label>OLD</label>
    <label>VIEIL</label>
  </arc>
  <arc  from="#T2"  to="#T2">
    <label>OLD</label>
    <label>VIEIL</label>
  </arc>
  <arc  from="#T3"  to="#T3">
    <label>OLD</label>
    <label>VIEUX</label>
  </arc>
  <arc  from="#T3"  to="#T5">
    <label>MEN</label>
    <label>HOMMES</label>
  </arc>
  <arc  from="#T4"  to="#T6">
    <label>COMES</label>
    <label>VIENT</label>
  </arc>
  <arc  from="#T5"  to="#T6">
    <label>COME</label>
    <label>VIENNENT</label>
  </arc>
</graph>
```

### 20.1.2 Family Trees

The next example provides an encoding a portion of a family tree[121] in which nodes are used to represent individuals and parents of individuals, and arcs are used to represent common parentage and descent links. Let us suppose, further, that information about individuals is contained in feature structures, which are contained in feature-structure libraries elsewhere (see *19.4. Feature Libraries and Feature-Value Libraries*). We can use the *value* attribute on <node> elements to point to those feature structures. In this particular representation of the graph, nodes representing females are framed by ovals, nodes representing males are framed by boxes, and nodes representing parents are framed by diamonds.

[121] The family tree is that of the mathematician and philosopher Bertrand Russell, whose third wife was commonly known as Peter. The information presented here is taken from Pereira and Shieber (1987).




```
<graph type="family_tree" order="13"
  size="12">
  <node xml:id="KATHR"
    value="http://example.com/russell-fs/tei/kr1" inDegree="0" outDegree="1">
    <label>Katherine</label>
  </node>
  <node xml:id="AMBER"
    value="http://example.com/russell-fs/tei/ar1" inDegree="0" outDegree="1">
    <label>Amberley</label>
  </node>
  <node xml:id="KAR" inDegree="2"
    outDegree="3">
    <label>K+A</label>
  </node>
  <node xml:id="BERTR"
    value="http://example.com/russell-fs/tei/br1" inDegree="1" outDegree="2">
    <label>Bertrand</label>
  </node>
  <node xml:id="PETER"
    value="http://example.com/russell-fs/tei/pr1" inDegree="0" outDegree="1">
    <label>Peter</label>
  </node>
  <node xml:id="DORAR"
    value="http://example.com/russell-fs/tei/dr1" inDegree="0" outDegree="1">
    <label>Dora</label>
  </node>
  <node xml:id="PBR" inDegree="2"
    outDegree="1">
    <label>P+B</label>
  </node>
  <node xml:id="DBR" inDegree="2"
    outDegree="2">
    <label>D+B</label>
  </node>
  <node xml:id="FRANR"
    value="http://example.com/russell-fs/tei/fr1" inDegree="1" outDegree="0">
    <label>Frank</label>
  </node>
  <node xml:id="RACHR"
    value="http://example.com/russell-fs/tei/rr1" inDegree="1" outDegree="0">
    <label>Rachel</label>
  </node>
```

```
  <node xml:id="CONRR"
    value="http://example.com/russell-fs/tei/cr1" inDegree="1" outDegree="0">
    <label>Conrad</label>
  </node>
  <node xml:id="KATER"
    value="http://example.com/russell-fs/tei/kr2" inDegree="1" outDegree="0">
    <label>Kate</label>
  </node>
  <node xml:id="JOHNR"
    value="http://example.com/russell-fs/tei/jr1" inDegree="1" outDegree="0">
    <label>John</label>
  </node>
  <arc from="#KATHR" to="#KAR">
    <label>Mo</label>
  </arc>
  <arc from="#AMBER" to="#KAR">
    <label>Fa</label>
  </arc>
  <arc from="#KAR" to="#BERTR">
    <label>So</label>
  </arc>
  <arc from="#KAR" to="#FRANR">
    <label>So</label>
  </arc>
  <arc from="#KAR" to="#RACHR">
    <label>Da</label>
  </arc>
  <arc from="#PETER" to="#PBR">
    <label>Mo</label>
  </arc>
  <arc from="#BERTR" to="#PBR">
    <label>Fa</label>
  </arc>
  <arc from="#PBR" to="#CONRR">
    <label>So</label>
  </arc>
  <arc from="#DORAR" to="#DBR">
    <label>Mo</label>
  </arc>
  <arc from="#BERTR" to="#DBR">
    <label>Fa</label>
  </arc>
  <arc from="#DBR" to="#KATER">
    <label>Da</label>
  </arc>
  <arc from="#DBR" to="#JOHNR">
    <label>So</label>
  </arc>
</graph>
```

### 20.1.3 Historical Interpretation

For our final example, we represent graphically the relationships among various geographic areas mentioned in a seventeenth-century Scottish document. The document itself is a ‘sasine’, which records a grant of land from the earl of Argyll to one Donald McNeill, and reads in part as follows (abbreviations have been expanded silently, and ‘[…]’ marks illegible passages):

> Item instrument of Sasine given the said Hector Mcneil confirmed and dated 28 May 1632 [...] at Edinburgh upon the 15 June 1632
>
> Item ane charter granted by Archibald late earl of Argyle and Donald McNeill of Gallachalzie wh makes mention that ... the said late Earl yields and grants to the said Donald MacNeill ...
>
> All and hail the two merk land of old extent of Gallachalzie with the pertinents by and in the lordship of Knapdale within the sherrifdome of Argyll
>
> [description of other lands granted follows ...]
>
> This Charter is dated at Inverary the 15th May 1669

In this example, we are concerned with the land and pertinents (i.e. accompanying sources of revenue) described as 'the two merk land of old extent of Gallachalzie with the pertinents by and in the lordship of Knapdale within the sherrifdom of Argyll'.

The passage concerns the following pieces of land:

- the Earl of Argyll's land (i.e. the lands granted by this clause of the sasine)
- two mark of land in Gallachalzie
- the pertinents for this land
- the Lordship of Knapdale
- the sherrifdom of Argyll

We will represent these geographic entities as nodes in a graph. Arcs in the graph will represent the following relationships among them:

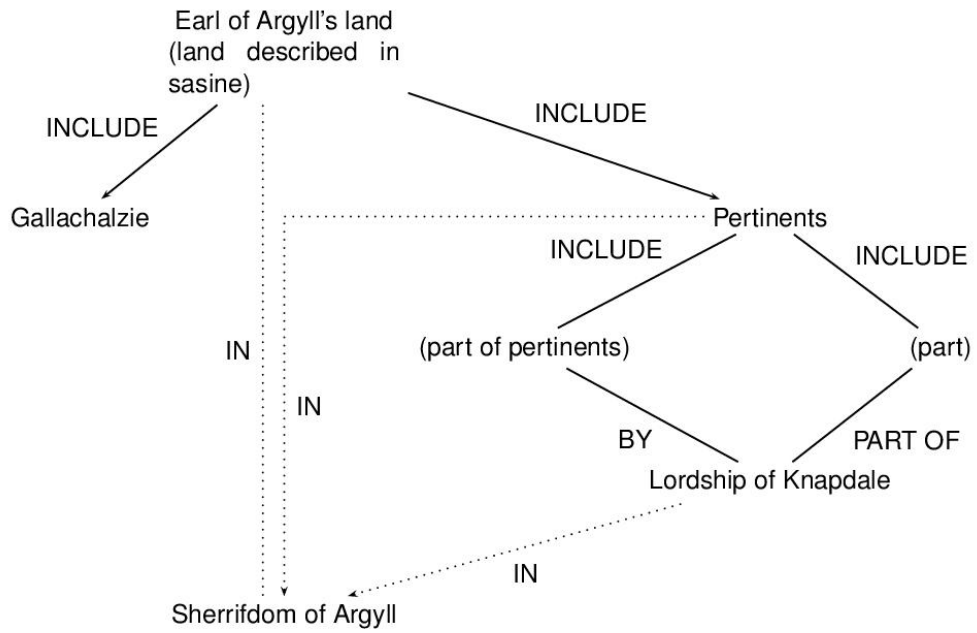
- containment (INCLUDE)
- location within (IN)
- contiguity (BY)
- constituency (PART OF)

Note that these relationships are logically related: 'include' and 'in', for example, are inverses of each other: the Earl of Argyll's land includes the parcel in Gallachalzie, and the parcel is therefore in the Earl of Argyll's land. Given an explicit set of inference rules, an appropriate application could use the graph we are constructing to infer the logical consequences of the relationships we identify.

Let us assume that feature-structure analyses are available which describe Gallachalzie, Knapdale, and Argyll. We will link to those feature structures using the *value* attribute on the nodes representing those places. However, there may be some uncertainty as to which noun phrase is modified by the phrase 'within the sheriffdome of Argyll': perhaps the entire lands (land and pertinents) are in Argyll, perhaps just the pertinents are, or perhaps only Knapdale is (together with the portion of the pertinents which is in Knapdale). We will represent all three of these interpretations in the graph; they are, however, mutually exclusive, which we represent using the *exclude* attribute defined in chapter *17. Linking, Segmentation, and Alignment.*[122]

We represent the graph and its encoding as follows, where the dotted lines in the graph indicate the mutually exclusive arcs; in the encoding, we use the *exclude* attribute to indicate those arcs.

[122] That is, the three syntactic interpretations of the clause are mutually exclusive. The notion that the pertinents are in Argyll is clearly not inconsistent with the notion that both the land in Gallachalzie and the pertinents are in Argyll. The graph given here describes the possible interpretations of the clause itself, not the sets of inferences derivable from each syntactic interpretation, for which it would be convenient to use the facilities described in chapter *19. Feature Structures.*




The graph formalizes the following relationships:

- the Earl of Argyll's land *includes* (the parcel of land in) Gallachalzie
- the Earl of Argyll's land *includes* the pertinents of that parcel
- the pertinents are (in part) *by* the Lordship of Knapdale
- the pertinents are (in part) *part of* the Lordship of Knapdale
- the Earl of Argyll's land, or the pertinents, or the Lordship of Knapdale, is *in* the Sherrifdom of Argyll

We encode the graph thus:

```
<graph type="directed" order="7" size="9">
  <node xml:id="EARL">
    <label>Earl of Argyll's land</label>
  </node>
  <node xml:id="GALL"
    value="http://example.com/people/scots#gall">
    <label>Gallachalzie</label>
  </node>
  <node xml:id="PERT">
    <label>Pertinents</label>
  </node>
  <node xml:id="PER1">
    <label>Pertinents part</label>
  </node>
  <node xml:id="PER2">
    <label>Pertinents part</label>
  </node>
  <node xml:id="KNAP"
    value="http://example.com/people/scots#knapfs">
    <label>Lordship of Knapdale</label>
  </node>
```

```
  <node  xml:id="ARGY"
    value="http://example.com/people/scots#argyfs">
    <label>Sherrifdome of Argyll</label>
  </node>
  <arc  xml:id="EARLGALL"  from="#EARL"
    to="#GALL">
    <label>INCLUDE</label>
  </arc>
  <arc  xml:id="EARLARGY"  from="#EARL"
    to="#ARGY"  exclude="#PERTARGY #KNAPARGY">
    <label>IN</label>
  </arc>
  <arc  xml:id="EARLPERT"  from="#EARL"
    to="#PERT">
    <label>INCLUDE</label>
  </arc>
  <arc  xml:id="PERTPER1"  from="#PERT"
    to="#PER1">
    <label>INCLUDE</label>
  </arc>
  <arc  xml:id="PERTPER2"  from="#PERT"
    to="#PER2">
    <label>INCLUDE</label>
  </arc>
  <arc  xml:id="PERTARGY"  from="#PERT"
    to="#ARGY"  exclude="#EARLARGY #KNAPARGY">
    <label>IN</label>
  </arc>
  <arc  xml:id="PER1KNAP"  from="#PER1"
    to="#KNAP">
    <label>BY</label>
  </arc>
  <arc  xml:id="PER2KNAP"  from="#PER2"
    to="#KNAP">
    <label>PART OF</label>
  </arc>
  <arc  xml:id="KNAPARGY"  from="#KNAP"
    to="#ARGY"  exclude="#EARLARGY #PERTARGY">
    <label>IN</label>
  </arc>
</graph>
```

## 20.2 Trees

A *tree* is a connected acyclic graph. That is, it is possible in a tree graph to follow a path from any vertex to any other vertex, but there are no paths that lead from any vertex to itself. A rooted tree is a directed graph based on a tree; that is, the arcs in the graph correspond to the arcs of a tree such that there is exactly one node, called the *root*, for which there is a path from that node to all other nodes in the graph. For our purposes, we may ignore all trees except for rooted trees, and hence we shall use the <tree> element for rooted trees, and the <root> element for its root. The nodes adjacent to a given node are called its *children*, and the node adjacent from a given node is called its *parent*. Nodes with both a parent and children are called *internal nodes*, for which we use the <iNode> element. A node with no children is tagged as a <leaf>. If the children of a node are ordered from left to right, then we say that that node is *ordered*. If all the nodes of a tree are ordered, then we say that the tree is an *ordered tree*. If some of the nodes of a tree are ordered and others are not, then the tree is a *partially ordered tree*. The ordering of nodes and trees may be specified by an attribute; we take the default ordering for trees to be ordered, that roots inherit their ordering from the trees in which they occur, and internal nodes inherit their ordering from their

parents. Finally, we permit a node to be specified as following other nodes, which (when its parent is ordered) it would be assumed to precede, giving rise to crossing arcs. The elements used for the encoding of trees have the following descriptions and attributes.

`<tree>` (tree) encodes a tree, which is made up of a root, internal nodes, leaves, and arcs from root to leaves.

- `@arity` gives the maximum number of children of the root and internal nodes of the tree.
- `@ord` (ordered) indicates whether or not the tree is ordered, or if it is partially ordered.
- `@order` gives the order of the tree, i.e., the number of its nodes.

`<root>` (root node) represents the root node of a tree.

- `@value` identifies the root node of the network by pointing to a feature structure or other analytic element.
- `@children` identifies the elements which are the children of the root node.
- `@ord` (ordered) indicates whether or not the root is ordered.
- `@outDegree` gives the out degree of the root, the number of its children.

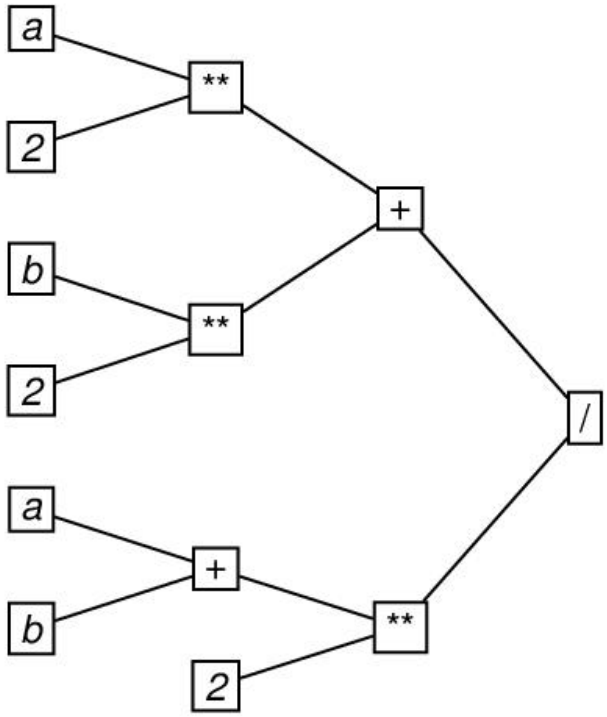
`<iNode>` (intermediate (or internal) node) represents an intermediate (or internal) node of a tree.

- `@value` indicates an intermediate node, which is a feature structure or other analytic element.
- `@children` provides a list of identifiers of the elements which are the children of the intermediate node.
- `@parent` provides the identifier of the element which is the parent of this node.
- `@ord` (ordered) indicates whether or not the internal node is ordered.
- `@follow` provides the identifier of an element which this node follows.
- `@outDegree` gives the out degree of an intermediate node, the number of its children.

`<leaf>` (leaf) encodes the leaves (terminal nodes) of a tree.

- `@value` provides a pointer to a feature structure or other analytic element.
- `@parent` provides the identifier of parent of a leaf.
- `@follow` provides an identifier of an element which this leaf follows.

Here is an example of a tree. It represents the order in which the operators of addition (symbolized by `+`), exponentiation (symbolized by `**`) and division (symbolized by `/`) are applied in evaluating the arithmetic formula `((a**2)+(b**2))/((a+b)**2)`. In drawing the graph, the root is placed on the far right, and directionality is presumed to be to the left.




```
<tree n="ex1" arity="2" ord="true" order="12">
  <root xml:id="G-DIV1"
    children="#PLU1 #EXP1">
    <label>/</label>
  </root>
  <iNode xml:id="PLU1" parent="#G-DIV1"
    children="#EXP2 #EXP3">
    <label>+</label>
  </iNode>
  <iNode xml:id="EXP1" parent="#G-DIV1"
    children="#PLU2 #NUM2.3">
    <label>**</label>
  </iNode>
  <iNode xml:id="EXP2" parent="#PLU1"
    children="#VARA1 #NUM2.1">
    <label>**</label>
  </iNode>
  <iNode xml:id="EXP3" parent="#PLU1"
    children="#VARB1 #NUM2.2">
    <label>**</label>
  </iNode>
  <iNode xml:id="PLU2" parent="#EXP1"
    children="#VARA2 #VARB2">
    <label>+</label>
  </iNode>
  <leaf xml:id="VARA1" parent="#EXP2">
    <label>a</label>
  </leaf>
  <leaf xml:id="NUM2.1" parent="#EXP2">
    <label>2</label>
  </leaf>
```

```
  <leaf  xml:id="VARB1"  parent="#EXP3">
    <label>b</label>
  </leaf>
  <leaf  xml:id="NUM2.2"  parent="#EXP3">
    <label>2</label>
  </leaf>
  <leaf  xml:id="VARA2"  parent="#PLU2">
    <label>a</label>
  </leaf>
  <leaf  xml:id="VARB2"  parent="#PLU2">
    <label>b</label>
  </leaf>
  <leaf  xml:id="NUM2.3"  parent="#EXP1">
    <label>2</label>
  </leaf>
</tree>
```

In this encoding, the *arity* attribute represents the *arity* of the tree, which is the greatest value of the *outDegree* attribute for any of the nodes in the tree. If, as in this case, *arity* is 2, we say that the tree is a *binary* tree.

Since the left-to-right (or top-to-bottom!) order of the children of the two + nodes does not affect the arithmetic result in this case, we could represent in this tree all of the arithmetically equivalent formulas involving its leaves, by specifying the attribute *ord* as false on those two <iNode> elements, the attribute *ord* as true on the <root> and other <iNode> elements, and the attribute *ord* as partial on the <tree> element, as follows.

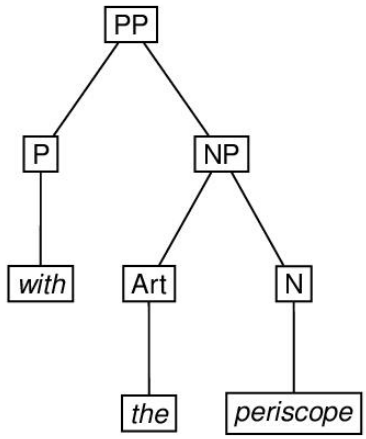
```
<tree  n="ex2"  ord="partial"  arity="2"
  order="13">
  <root  xml:id="divi1"  ord="true"
    children="#plu1 #exp1">
    <label>/</label>
  </root>
  <iNode  xml:id="plu1"  ord="false"
    parent="#divi1"  children="#exp2 #exp3">
    <label>+</label>
  </iNode>
  <iNode  xml:id="exp1"  ord="true"
    parent="#divi1"  children="#plu2 #num2.3">
    <label>**</label>
  </iNode>
  <iNode  xml:id="exp2"  ord="true"
    parent="#plu1"  children="#vara1 #num2.1">
    <label>**</label>
  </iNode>
  <iNode  xml:id="exp3"  ord="true"
    parent="#plu1"  children="#varb1 #num2.2">
    <label>**</label>
  </iNode>
  <iNode  xml:id="plu2"  ord="false"
    parent="#exp1"  children="#vara2 #varb2">
    <label>+</label>
  </iNode>
  <leaf  xml:id="vara1"  parent="#exp2">
    <label>a</label>
  </leaf>
  <leaf  xml:id="num2.1"  parent="#exp2">
    <label>2</label>
  </leaf>
```

```
  <leaf  xml:id="varb1"  parent="#exp3">
    <label>b</label>
  </leaf>
  <leaf  xml:id="num2.2"  parent="#exp3">
    <label>2</label>
  </leaf>
  <leaf  xml:id="vara2"  parent="#plu2">
    <label>a</label>
  </leaf>
  <leaf  xml:id="varb2"  parent="#plu2">
    <label>b</label>
  </leaf>
  <leaf  xml:id="num2.3"  parent="#exp1">
    <label>2</label>
  </leaf>
</tree>
```

This encoding represents all of the following:

- `((a**2)+(b**2))/((a+b)**2)`
- `((b**2)+(a**2))/((a+b)**2)`
- `((a**2)+(b**2))/((b+a)**2)`
- `((b**2)+(a**2))/((a+b)**2)`

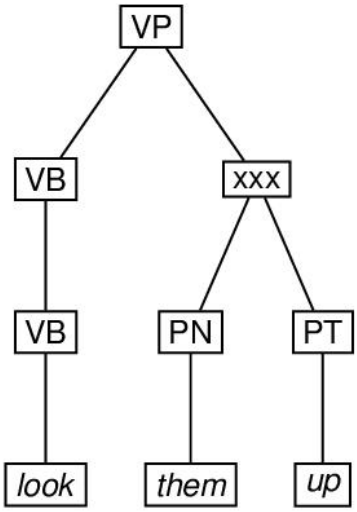
Linguistic phrase structure is very commonly represented by trees. Here is an example of phrase structure represented by an ordered tree with its root at the top, and a possible encoding.




```
<tree  n="ex3"  ord="true"  arity="2"  order="8">
  <root  xml:id="GD-PP1"
    children="#GD-P1 #GD-NP1">
    <label>PP</label>
  </root>
  <iNode  xml:id="GD-P1"  parent="#GD-PP1"
    children="#GD-WITH1">
    <label>P</label>
  </iNode>
  <leaf  xml:id="GD-WITH1"  parent="#GD-P1">
    <label>with</label>
  </leaf>
  <iNode  xml:id="GD-NP1"  parent="#GD-PP1"
```

```
    children="#GD-THE1 #GD-PERI1">
    <label>NP</label>
  </iNode>
  <iNode  xml:id="GD-ART1"  parent="#GD-NP1"
    children="#GD-THE1">
    <label>Art</label>
  </iNode>
  <leaf  xml:id="GD-THE1"  parent="#GD-ART1">
    <label>the</label>
  </leaf>
  <iNode  xml:id="GD-N1"  parent="#GD-NP1"
    children="#GD-PERI1">
    <label>N</label>
  </iNode>
  <leaf  xml:id="GD-PERI1"  parent="#GD-N1">
    <label>periscope</label>
  </leaf>
</tree>
```

Finally, here is an example of an ordered tree, in which a particular node which ordinarily would precede another is specified as following it. In the drawing, the xxx symbol indicates that the arc from VB to PT crosses the arc from VP to PN.




```
<tree  n="ex4"  arity="2"  order="8"  ord="true">
  <leaf  xml:id="GD-LOOK1"  parent="#GD-VB2">
    <label>look</label>
  </leaf>
  <leaf  xml:id="GD-THEM1"  parent="#GD-PN1">
    <label>them</label>
  </leaf>
  <leaf  xml:id="GD-UP1"  parent="#GD-PT1">
    <label>up</label>
  </leaf>
  <iNode  xml:id="GD-VB2"  parent="#GD-VB1"
    children="#GD-LOOK1">
    <label>VB</label>
  </iNode>
  <iNode  xml:id="GD-PN1"  parent="#GD-VP1"
```

```
    children="#GD-THEM1">
    <label>PN</label>
  </iNode>
  <iNode  xml:id="GD-PT1"  parent="#GD-VB1"
    children="#GD-UP1"  follow="#GD-PN1">
    <label>PT</label>
  </iNode>
  <iNode  xml:id="GD-VB1"  parent="#GD-VP1"
    children="#GD-VB2 #GD-PT1">
    <label>VB</label>
  </iNode>
  <root  xml:id="GD-VP1"
    children="#GD-VB1 #GD-PN1">
    <label>VP</label>
  </root>
</tree>
```

## 20.3 Another Tree Notation

In this section, we present an alternative to the method of representing the structure of ordered rooted trees given in section *20.2. Trees*, which is based on the observation that any node of such a tree can be thought of as the root of the subtree that it dominates. Thus subtrees can be thought of as the same type as the trees they are embedded in, hence the designation <eTree>, for *embedding tree*. Whereas in a <tree> the relationship among the parts is indicated by the *children* attribute, and by the names of the elements <root>, <iNode>, and <leaf>, the relationship among the parts of an <eTree> is indicated simply by the arrangement of their content. However, we have chosen to enable encoders to distinguish the terminal elements of an <eTree> by means of the empty <eLeaf> element, though its use is not required; the <eTree> element can also be used to identify the terminal nodes of <eTree> elements. We also provide a <triangle> element, which can be thought of as an underspecified <eTree>, i.e. an <eTree> in which certain information has been left out. In addition, we provide a <forest> element, which consists of one or more <tree>, <eTree>, or <triangle> elements, and a <listForest> element, which consists of one or more <forest> elements. The elements used for the encoding of embedding trees and the units containing them have the following descriptions and attributes.

`<eTree>` (embedding tree) provides an alternative to the <tree> element for representing ordered rooted tree structures.

- `@value` provides the value of an embedding tree, which is a feature structure or other analytic element.

`<triangle>` (underspecified embedding tree, so called because of its characteristic shape when drawn) provides for an underspecified <eTree>, that is, an <eTree> with information left out.

- `@value` supplies a value for the triangle, in the form of the identifier of a feature structure or other analytic element.

`<eLeaf>` (leaf or terminal node of an embedding tree) provides explicitly for a leaf of an embedding tree, which may also be encoded with the <eTree> element.

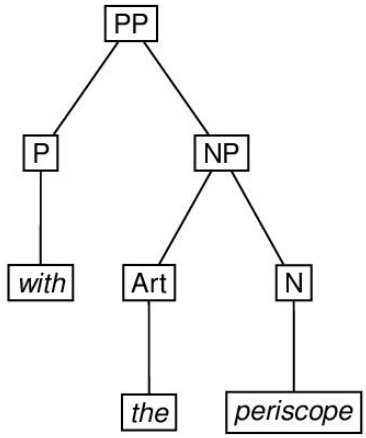
- `@value` indicates the value of an embedding leaf, which is a feature structure or other analytic element.

`<forest>` (forest) provides for groups of rooted trees.

`<listForest>` provides for lists of forests.

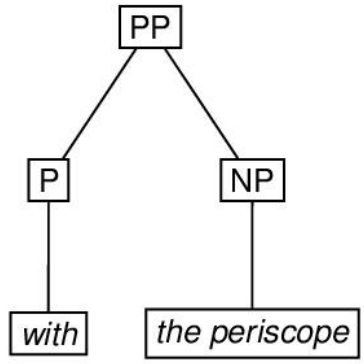
Like the <root>, <iNode>, and <leaf> of a <tree>, the <eTree>, <triangle> and <eLeaf> elements may also have *value* attributes and <label> children.

To illustrate the use of the <eTree> and <eLeaf> elements, here is an encoding of the second example in section *20.2. Trees*, repeated here for convenience.



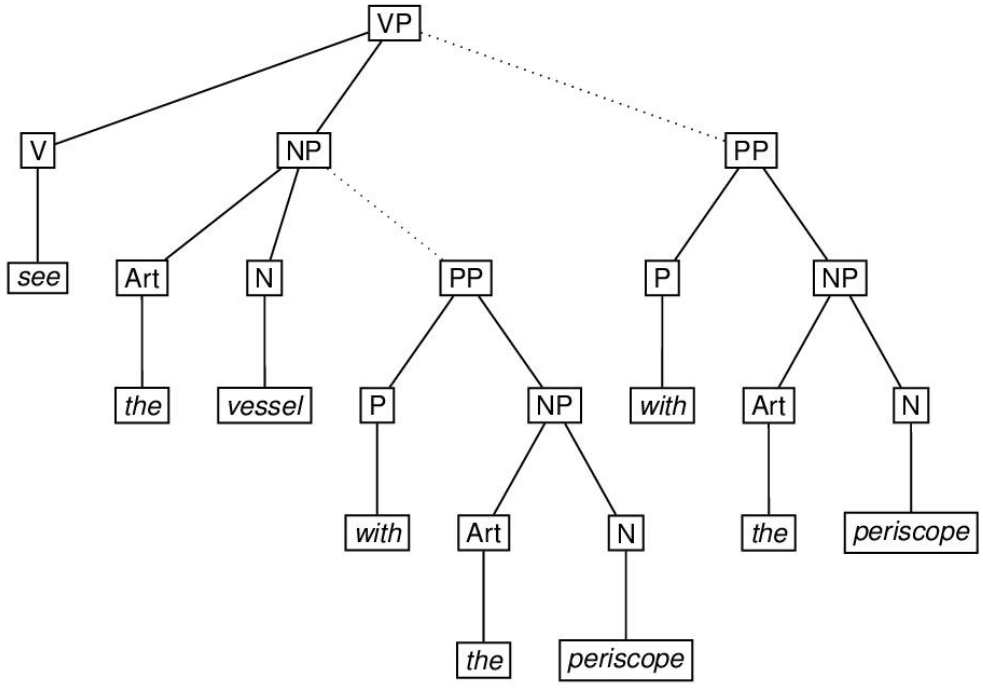

```
<eTree  n="ex1">
  <label>PP</label>
  <eTree>
    <label>P</label>
    <eLeaf>
      <label>with</label>
    </eLeaf>
  </eTree>
  <eTree>
    <label>NP</label>
    <eTree>
      <label>Art</label>
      <eLeaf>
        <label>the</label>
      </eLeaf>
    </eTree>
    <eTree>
      <label>N</label>
      <eLeaf>
        <label>periscope</label>
      </eLeaf>
    </eTree>
  </eTree>
</eTree>
```

Next, we provide an encoding, using the <triangle> element, in which the internal structure of the <eTree> labeled NP is omitted.



```
<eTree  n="ex2">
  <label>PP</label>
  <eTree>
    <label>P</label>
    <eLeaf>
      <label>with</label>
    </eLeaf>
  </eTree>
  <triangle>
    <label>NP</label>
    <eLeaf>
      <label>the periscope</label>
    </eLeaf>
  </triangle>
</eTree>
```

Ambiguity involving alternative tree structures associated with the same terminal sequence can be encoded relatively conveniently using a combination of the *exclude* and *copyOf* attributes described in sections *17.8. Alternation* and *17.6. Identical Elements and Virtual Copies*. In the simplest case, an <eTree> may be part of the content of exactly one of two different <eTree> elements. To mark it up, the embedded <eTree> may be fully specified within one of the embedding <eTree> elements to which it may belong, and a virtual copy, specified by the *copyOf* attribute, may appear on the other. In addition, each of the embedded elements in question is specified as excluding the other, using the *exclude* attribute. To illustrate, consider the English phrase *see the vessel with the periscope*, which may be considered to be structurally ambiguous, depending on whether the phrase *with the periscope* is a modifier of the phrase *the vessel* or a modifier of the phrase *see the vessel*. This ambiguity is indicated in the sketch of the ambiguous tree by means of the dotted-line arcs. The markup using the *copyOf* and *exclude* attributes follows the sketch.




```
<eTree  n="ex3">
  <label>VP</label>
  <eTree>
    <label>V</label>
```

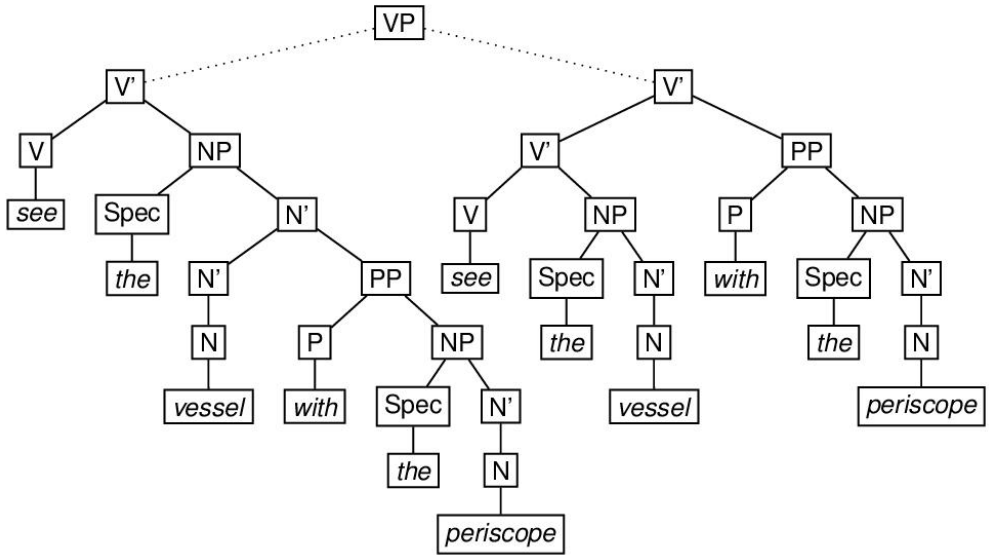
```
    <eLeaf>
      <label>see</label>
    </eLeaf>
  </eTree>
  <eTree>
    <label>NP</label>
    <eTree>
      <label>Art</label>
      <eLeaf>
        <label>the</label>
      </eLeaf>
    </eTree>
    <eTree>
      <label>N</label>
      <eLeaf>
        <label>vessel</label>
      </eLeaf>
    </eTree>
    <eTree  xml:id="GD-PPA"  exclude="#GD-PPB">
      <label>PP</label>
      <eTree>
        <label>P</label>
        <eLeaf>
          <label>with</label>
        </eLeaf>
      </eTree>
      <eTree>
        <label>NP</label>
        <eTree>
          <label>Art</label>
          <eLeaf>
            <label>the</label>
          </eLeaf>
        </eTree>
        <eTree>
          <label>N</label>
          <eLeaf>
            <label>periscope</label>
          </eLeaf>
        </eTree>
      </eTree>
    </eTree>
  </eTree>
  <eTree  xml:id="GD-PPB"  copyOf="#GD-PPA"
    exclude="#GD-PPA">
    <label>PP</label>
  </eTree>
</eTree>
```

To indicate that one of the alternatives is selected, one may specify the *select* attribute on the highest <eTree> as either #GD-PPA or #GD-PPB; see section *17.8. Alternation*.

Depending on the grammar one uses to associate structures with examples like *see the man with the periscope*, the representations may be more complicated than this. For example, adopting a version of the *X-bar* theory of phrase structure originated by Jackendoff,[123] the attachment of a modifier may require the creation of an intermediate node

[123] ?

which is not required when the attachment is not made, as shown in the following diagram. A possible encoding of this ambiguous structure immediately follows the diagram.

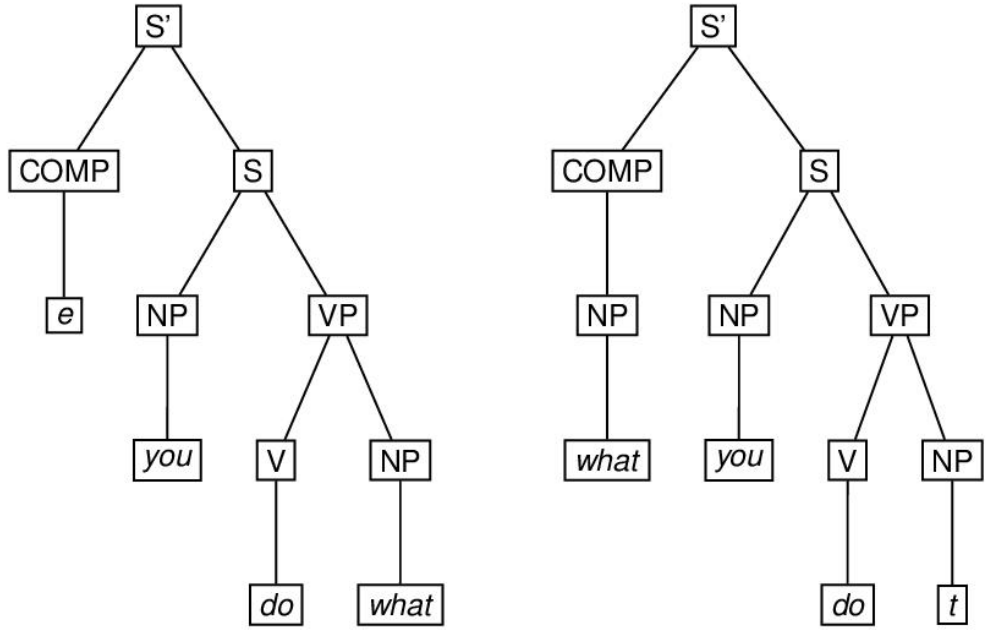



```
<eTree  n="ex4">
  <label>VP</label>
  <eTree  xml:id="VBARA"  exclude="#VBARB">
    <label>V'</label>
    <eTree  xml:id="VA">
      <label>V</label>
      <eLeaf>
        <label>see</label>
      </eLeaf>
    </eTree>
    <eTree>
      <label>NP</label>
      <eTree  xml:id="SPEC1A">
        <label>Spec</label>
        <eLeaf>
          <label>the</label>
        </eLeaf>
      </eTree>
      <eTree>
        <label>N'</label>
        <eTree  xml:id="NBAR2A">
          <label>N'</label>
          <eTree>
            <label>N</label>
            <eLeaf>
              <label>vessel</label>
            </eLeaf>
          </eTree>
        </eTree>
        <eTree  xml:id="PPA1">
          <label>PP</label>
          <eTree>
            <label>P</label>
            <eLeaf>
              <label>with</label>
```

```
          </eLeaf>
        </eTree>
        <eTree>
          <label>NP</label>
          <eTree>
            <label>Spec</label>
            <eLeaf>
              <label>the</label>
            </eLeaf>
          </eTree>
          <eTree>
            <label>N'</label>
            <eTree>
              <label>N</label>
              <eLeaf>
                <label>periscope</label>
              </eLeaf>
            </eTree>
          </eTree>
        </eTree>
      </eTree>
    </eTree>
  </eTree>
  <eTree xml:id="VBARB" exclude="#VBARA">
    <label>V'</label>
    <eTree>
      <label>V'</label>
      <eTree xml:id="VB" copyOf="#VA">
        <label>V</label>
      </eTree>
      <eTree>
        <label>NP</label>
        <eTree xml:id="SPEC1B" copyOf="#SPEC1A">
          <label>Spec</label>
        </eTree>
        <eTree xml:id="NBAR2B" copyOf="#NBAR2A">
          <label>N'</label>
        </eTree>
      </eTree>
    </eTree>
    <eTree xml:id="PPB" copyOf="#PPA1">
      <label>PP</label>
    </eTree>
  </eTree>
</eTree>
```

A *derivation* in a generative grammar is often thought of as a set of trees. To encode such a derivation, one may use the <forest> element, in which the trees may be marked up using the <tree>, the <eTree>, or the <triangle> element. The *type* attribute may be used to specify what kind of derivation it is. Here is an example of a two-tree forest, involving application of the 'wh-movement' transformation in the derivation of *what you do* (as in *this is what you do*) from the underlying *you do what*.[124]

[124]The symbols `e` and `t` denote special theoretical constructs (*empty category* and *trace* respectively), which need not concern us here.




```
<forest  n="ex5"  type="derivation-syntactic">
  <eTree  n="Stage 1"  xml:id="S1SBAR">
    <label>S'</label>
    <eTree  xml:id="S1COMP">
      <label>COMP</label>
      <eLeaf  xml:id="S1E">
        <label>e</label>
      </eLeaf>
    </eTree>
    <eTree  xml:id="S1S">
      <label>S</label>
      <eTree  xml:id="S1NP1">
        <label>NP</label>
        <eLeaf>
          <label>you</label>
        </eLeaf>
      </eTree>
      <eTree  xml:id="S1VP">
        <label>VP</label>
        <eTree  xml:id="S1V">
          <label>V</label>
          <eLeaf>
            <label>do</label>
          </eLeaf>
        </eTree>
        <eTree  xml:id="S1NP2">
          <label>NP</label>
          <eLeaf  xml:id="S1WH">
            <label>what</label>
          </eLeaf>
        </eTree>
      </eTree>
    </eTree>
  </eTree>
  <eTree  n="Stage 2"  xml:id="S2SBAR"
```

```
    corresp="#S1SBAR">
    <label>S'</label>
    <eTree xml:id="S2COMP" corresp="#S1COMP">
      <label>COMP</label>
      <eTree copyOf="#S1NP2" corresp="#S1E">
        <label>NP</label>
      </eTree>
    </eTree>
    <eTree xml:id="S2S" corresp="#S1S">
      <label>S</label>
      <eTree xml:id="S2NP1" copyOf="#S1NP1">
        <label>NP</label>
      </eTree>
      <eTree xml:id="S2VP" corresp="#S1VP">
        <label>VP</label>
        <eTree xml:id="S2V" copyOf="#S1V">
          <label>V</label>
        </eTree>
        <eTree xml:id="S2NP2" corresp="#S1NP2">
          <label>NP</label>
          <eLeaf corresp="#S1WH">
            <label>t</label>
          </eLeaf>
        </eTree>
      </eTree>
    </eTree>
  </eTree>
</forest>
```

In this markup, we have used *copyOf* attributes to provide virtual copies of elements in the tree representing the second stage of the derivation that also occur in the first stage, and the *corresp* attribute (see section *17.5. Correspondence and Alignment*) to link those elements in the second stage with corresponding elements in the first stage that are not copies of them.

If a group of forests (e.g. a full grammatical derivation including syntactic, semantic, and phonological subderivations) is to be articulated, the grouping element <listForest> may be used.

## 20.4 Representing Textual Transmission

A *stemma codicum* (sometimes called just *stemma*) is a tree-like graphic structure that has become traditional in manuscript studies for representing textual transmission. Consider the following hypothetical stemma:

The nodes in this stemma represent manuscripts; each has a label (a letter) which identifies it and also distinguishes whether the manuscript is extant, lost, or hypothetical. Extant manuscripts are identified by uppercase Latin letters or words beginning with uppercase Latin letters, e.g., L, shown as aqua in this example; manuscripts no longer existing, but providing readings which are attested e.g. by note or copy made before their disappearance, are identified by lowercase Latin letters, e.g., t, shown as magenta in this example; hypothetical stages in the textual transmission, which do not necessarily correspond to real manuscripts, are given lowercase Greek letters, e.g., α and shown as gold in this example. The stemma shown above thus suggests that (on the basis of similarities in the readings of the extant and lost manuscripts) L and t share textual material that is not shared with other manuscripts (represented in this case by δ) even though no physical manuscript attesting this stage in the textual transmission has ever been identified.

Manuscripts are copied from other manuscripts. The preceding stemma represents the hypothesis that all manuscripts go back to a common ancestor (α), that the tradition split after that stage into two (β and γ), etc. Descent by copying is indicated with a solid line. According to this model, α is the earliest common hypothetical stage that can be reconstructed, and all nodes below α have a single parent, that is, were copied from a single other stage in the tradition.

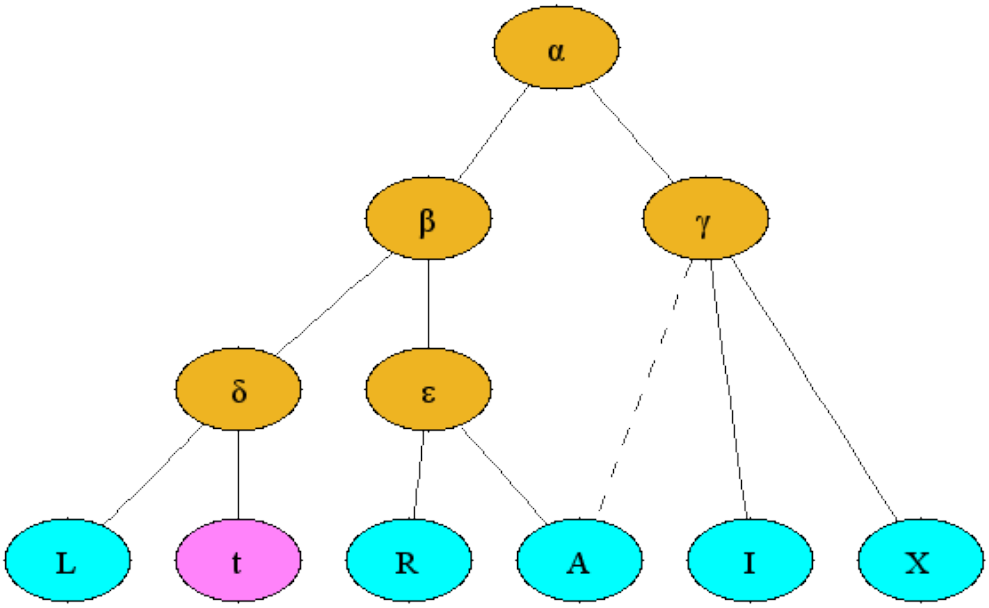



Figure 34: Example stemma

This familiar tree model is complicated because manuscripts sometimes show the influence of more than one ancestor. They may have been produced by a scribe who checked the text in one manuscript of the same work whilst copying from another, or perhaps made changes from his memory of a slightly different version of the text that he had read elsewhere. Alternatively, perhaps scribe A copied a manuscript from one source, scribe B made changes in it in the margins or between the lines (either by consulting another source directly or from memory), and another scribe then copied that manuscript, incorporating the changes into the body. Whatever the specific scenario, it is not uncommon for a manuscript to be based primarily on one source, but to incorporate features of another branch of the tradition. This mixed result is called *contamination*, and it is reflected in a stemma by a dotted line. Thus, the example above asserts that A is copied within the ε tradition, but is also contaminated from the γ tradition.

The utility of a stemma as a visualization tool is inversely proportional to the degree of contamination in the manuscript tradition. A tradition completely without contamination (called a *closed tradition*) yields a classic tree, easily represented graphically by a stemma. An *open tradition*, with substantial contamination, yields a spaghetti-like stemma characterized by crossing dotted lines, which is both difficult to read and not very informative.

The <eTree> element introduced in this chapter can be used to represent a closed tradition in a straightforward manner. Each non-terminal node is represented by a typed <eTree> element and each terminal node by an <eLeaf>. A <label> element provides a way of identifying each node, complementary to the global attributes *n* and *xml:id* attributes. For example, the closed part of the tradition headed by the label δ may be encoded as follows:

```
<eTree type="hypothetical">
  <label>δ</label>
  <eLeaf type="extant">
    <label>L</label>
  </eLeaf>
  <eLeaf type="lost">
    <label>t</label>
  </eLeaf>
</eTree>
```

To complete this representation, we need to show that the node labelled A is not derived solely from its parent node (labelled ε) but also demonstrates contamination from the node labelled γ. The easiest way to accomplish this is to include an appropriately-typed <ptr> element within the node in question, the *target* of which points to the

node labelled γ. This requires that this latter node be supplied with a value for its *xml:id* attribute. The complete representation is thus:

```
<eTree type="hypothetical">
  <label>α</label>
  <eTree type="hypothetical">
    <label>β</label>
    <eTree type="hypothetical">
      <label>δ</label>
      <eLeaf type="extant">
        <label>L</label>
      </eLeaf>
      <eLeaf type="lost">
        <label>t</label>
      </eLeaf>
    </eTree>
    <eTree type="hypothetical">
      <label>ε</label>
      <eLeaf type="extant">
        <label>R</label>
      </eLeaf>
      <eLeaf type="extant">
        <label>A</label>
        <ptr type="contamination"
          target="#gamma"/>
      </eLeaf>
    </eTree>
  </eTree>
  <eTree xml:id="gamma" type="hypothetical">
    <label>γ</label>
    <eLeaf type="extant">
      <label>I</label>
    </eLeaf>
    <eLeaf type="extant">
      <label>X</label>
    </eLeaf>
  </eTree>
</eTree>
```

In any substantial codicological project, it is likely that significantly more data will be required about the individual witnesses than indicated in the simple structures above. These Guidelines provide a rich variety of additional elements for representing such information: see in particular chapters *11. Manuscript Description*, *12. Representation of Primary Sources*, and *13. Critical Apparatus*.

## 20.5 Module for Graphs, Networks, and Trees

The module described in this chapter makes available the following components:

**Module nets: Graphs, networks, and trees**

- *Elements defined*: arc eLeaf eTree forest graph iNode leaf listForest node root tree triangle

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 21 Non-hierarchical Structures

XML employs a strongly hierarchical document model. At various points, these Guidelines discuss problems that arise when using XML to encode textual features that either do not naturally lend themselves to representation in a strictly hierarchical form or conflict with other hierarchies represented in the markup. Examples of such situations include:

- Conflict between the hierarchy established by the physical structure of a document (e.g., volume, page, column, line) and its rhetorical or linguistic structure (e.g., chapters, paragraphs, sentences, acts, scenes, etc.)
- Conflict between a verse text's metrical structure (e.g., its arrangement in stanzas and metrical lines) and its rhetorical or linguistic structure (e.g., phrases, sentences, and, for plays, acts, scenes, and speeches).
- Conflict between metrical, rhetorical, or linguistic structure and the representation of direct speech, especially if the quoted speech is interrupted by other elements (e.g., 'What', she asked, 'was that all about') or crosses metrical, rhetorical, or linguistic boundaries.
- Conflict between different analytical views or descriptions of a text or document, e.g., markup intended to encode diplomatic information about a word's appearance in a manuscript with markup intended to describe its morphology or pronunciation.

Non-nesting information poses fundamental problems for any XML-based encoding scheme, and it must be stated at the outset that no current solution combines all the desirable attributes of formal simplicity, capacity to represent all occurring or imaginable kinds of structures, suitability for formal or mechanical validation. The representation of non-hierarchical information is thus necessarily a matter of trade-offs among various sets of advantages and disadvantages.

These Guidelines support several methods for handling non-hierarchical information:

- redundant encoding of information in multiple forms (discussed in *21.1. Multiple Encodings of the Same Information*)
- the use of empty elements to delimit the boundaries of a non-nesting structure (discussed in *21.2. Boundary Marking with Empty Elements*)
- the division of a logically single non-nesting element into segments that nest properly in their immediate hierarchical context but can also be reconstituted virtually across these hierarchical boundaries (discussed *21.3. Fragmentation and Reconstitution of Virtual Elements*)
- stand-off markup: the annotation of information by pointing at it, rather than by placing XML tags within it (discussed in *21.4. Stand-off Markup*)

Some of these methods can be used in TEI-conformant documents. Others require extension.

In the sections which follow these techniques are described and their advantages and disadvantages are briefly discussed. The various solutions to the problem will be exemplified using extracts from two poems. The first is the opening quatrain from William Wordsworth's 'Scorn not the sonnet':

> Scorn not the sonnet; critic, you have frowned,
> Mindless of its just honours; with this key
> Shakespeare unlocked his heart; the melody
>
> Of this small lute gave ease to Petrarch's wound.

The second example is the third stanza from the fourth section of Robert Pinsky's 'Essay on Psychiatrists':

> Catholic woman of twenty-seven with five children
> And a first-rate body—pointed her finger
> at the back of one certain man and asked me,
> "Is that guy a psychiatrist?" and by god he was! "Yes,"
> She said, "He *looks* like a psychiatrist."
>
> Grown quiet, I looked at his pink back, and thought.

These two texts can be analysed in various ways. The first, which we might describe as the 'Metrical View', encodes the text according to its metrical features: line divisions (as here), stanzas or cantos in larger poems, and perhaps

prosodic features like stress or syllable patterns, alliteration, or rhyme. A second view, which we might describe as the 'Grammatical', encodes linguistic and rhetorical features: phonemes, morphemes, words, phrases, clauses, and sentences. A third view, the 'Dialogic', might concentrate on narrative voice: distinguishing between the narrator and their interlocutors and identifying individual segments as direct quotations. In our examples, we will restrict ourselves to relatively simple conflicts: for the *Metrical View* we will encode only metrical lines and line groups; for the *Grammatical View* we will restrict ourselves to encoding sentences; and for the *Dialogic View*, we only will distinguish direct quotation from other narration.

## 21.1 Multiple Encodings of the Same Information

Conceptually, the simplest method of disentangling two (or more) conflicting hierarchical views of the same information is to encode it twice (or more), each time capturing a single view.

Thus, for example, the *Metrical View* of 'Scorn not the sonnet' might be encoded as follows, using the <l> element to encode each metrical line:

```
<l>Scorn not the sonnet; critic, you have frowned,</l>
<l>Mindless of its just honours; with this key</l>
<l>Shakespeare unlocked his heart; the melody</l>
<l>Of this small lute gave ease to Petrarch's wound.</l>
```

The *Grammatical View* would be encoded by taking the same text and replacing the metrical markup with information about its sentence structure:

```
<p>
  <seg>Scorn not the sonnet;</seg>
  <seg>critic, you have frowned, Mindless of its just honours;</seg>
  <seg>with this key Shakespeare unlocked his heart;</seg>
  <seg>the melody Of this small lute gave ease to Petrarch's wound.</seg>
</p>
```

Likewise, the more complex passage from Pinsky could be encoded in three different ways to reflect the different metrical, grammatical, and dialogic views of its text:

```
<lg>
  <l>Catholic woman of twenty-seven with five children</l>
  <l>And a first-rate body—pointed her finger</l>
  <l>at the back of one certain man and asked me,</l>
  <l>"Is that guy a psychiatrist?" and by god he was! "Yes,"</l>
  <l>She said, "He <emph>looks</emph> like a psychiatrist."</l>
  <l>Grown quiet, I looked at his pink back, and thought.</l>
</lg>
```

```
<p>
  <seg>Catholic woman of twenty-seven with five children And a
     first-rate body—pointed her finger at the back of one certain man and
     asked me, "Is that guy a psychiatrist?" and by god he was!</seg>
</p>
<p>
  <seg>"Yes," She said, "He <emph>looks</emph> like a
     psychiatrist."</seg>
</p>
<p>
  <seg>Grown quiet, I looked at his pink back, and thought.</seg>
</p>
```

```
<ab>Catholic woman of twenty-seven with five children And a first-rate
body—pointed her finger at the back of one certain man and asked me,
<said>Is that guy a psychiatrist?</said> and by god he was!
<said>Yes,</said> She said, <said>He <emph>looks</emph> like a
    psychiatrist.</said> Grown quiet, I looked at his pink back, and
thought.</ab>
```

This method is TEI-conformant. Its advantages are that each way of looking at the information is explicitly represented in the data and that the individual views are simple to process. The disadvantages are that the method requires the maintenance of multiple copies of identical textual content (an invitation to inconsistency) and that there is no explicit indication that the various views, which might be in separate files, are related to each other: it might prove difficult to combine the views or access information from one view while processing the file that contains the encoding of another.[125]

## 21.2 Boundary Marking with Empty Elements

A second method for accommodating non-hierarchical objects in an XML document involves marking the start and end points of the non-nesting material. This prevents textual features that fall outside the privileged hierarchy from invalidating the document while identifying their beginnings and ends for further processing. The disadvantage of this method is that no single XML element represents the non-nesting material and, as a result, processing with XML technologies is significantly more difficult.

The empty elements used at each end are called *segment-boundary elements* or *segment-boundary delimiters*. There are several variations on this method of encoding.

For some common structural features, the TEI provides milestone elements that can be used to mark the beginning of a textual feature. These include <lb>, <pb>, <cb>, <handShift>, and the generic <milestone>. Using <lb>, for example, it is possible to indicate both the physical lineation of a poem on the page and its grammatical division into sentences:

```
<p>
  <seg>
    <lb  n="1"/>Scorn not the sonnet;</seg>; <seg>critic, you have
     frowned, <lb  n="2"/>Mindless of its just honours;</seg>
  <seg>with this
     key <lb  n="3"/>Shakespeare unlocked his heart;</seg>
  <seg>the melody
  <lb  n="4"/>Of this small lute gave ease to Petrarch's
     wound.</seg>
</p>
```

The use of these elements is by definition TEI-conformant. Care should be taken, however, that the meaning of the milestone elements is preserved: semantically, for example, <lb> is used to mark the start of a new (typographical) line. While in much modern poetry, typographical and metrical line divisions correspond, <lb> does not itself make a metrical claim: in encoding verse from sources, such as Old English manuscripts, where physical line beginnings are not used to indicate metrical lineation, the correspondence would break down entirely.

The segment boundaries also may be delimited by the generic <anchor> element. Attributes can then be used to indicate the type of feature being delimited and whether a given instance opens or closes the feature.

```
<l>
  <anchor  subtype="sentenceStart"
    type="delimiter"/>
```

[125] It has been shown, however, that it is possible to relate the different annotations in an indirect way: if the textual content of the annotations is identical, the very text can serve as a means for linking the different annotations, as described in ?.

```
Scorn not the sonnet;
<anchor  subtype="sentenceEnd"
   type="delimiter"/>
  <anchor  subtype="sentenceStart"
   type="delimiter"/> critic, you have frowned,
</l>
<l>Mindless of its just honours; <anchor  subtype="sentenceEnd"
   type="delimiter"/>
  <anchor  subtype="sentenceStart"
   type="delimiter"/> with this key</l>
<l>Shakespeare unlocked his heart; <anchor  subtype="sentenceEnd"
   type="delimiter"/>
  <anchor  subtype="sentenceStart"
   type="delimiter"/> the melody</l>
<l>Of this small lute gave ease to Petrarch's wound. <anchor  subtype="sentenceEnd"
   type="delimiter"/>
</l>
```

This method is TEI-conformant.

Another approach is to design custom elements that provide richer information about the feature being delimited or its boundaries. This information can be included as attribute values or as part of the element name itself: e.g., <boundaryStart element="sentence"/>… <boundaryEnd element="sentence"/>, <sentenceBoundary position="start"/>… <sentenceBoundary position="end"/>, or <sentenceBoundaryStart/>… <sentenceBoundaryEnd/>:

```
<l
   xmlns:n="http://www.example.org/ns/nonTEI">
  <n:sentenceBoundaryStart/>Scorn not the sonnet;
<n:sentenceBoundaryEnd/>
  <n:sentenceBoundaryStart/>critic, you have frowned,
</l>
<l>Mindless of its just honours; <n:sentenceBoundaryEnd/>
  <n:sentenceBoundaryStart/>with this key</l>
<l>Shakespeare unlocked his heart; <n:sentenceBoundaryEnd/>
  <n:sentenceBoundaryStart/>the melody</l>
<l>Of this small lute gave ease to Petrarch's wound. <n:sentenceBoundaryEnd/>
</l>
```

If the custom elements can be replaced by TEI elements and attributes without loss of information, this method is TEI-conformant (see *24.4. Conformance*); if the custom elements introduce information or distinctions that cannot be captured using standard TEI elements, the method is an extension.

Finally, elements that are normally used to encode nesting textual features (e.g., <said>, <seg>, <l>, etc.) can be adapted so that they serve as empty segment boundary delimiters when the features they encode cross-hierarchical boundaries. Additional attributes (*sID* and *eID* in the example below) are added to these elements in order to allow the unambiguous correlation of start and end points. This method has been introduced in the markup literature under various names, including Trojan milestones, HORSE markup, CLIX, and COLT. It is described in detail by ?):

```
<lg
   xmlns:hr="http://www.example.org/ns/nonTEI">
  <l>
   <seg>Scorn not the sonnet;</seg>
   <hr:s  sID="s02"/>critic, you have frowned, </l>
  <l>Mindless of its just honours; <hr:s  eID="s02"/>
   <hr:s  sID="s03"/>with this key </l>
```

```
  <l>Shakespeare unlocked his heart; <hr:s  eID="s03"/>
    <hr:s  sID="s04"/>the melody </l>
  <l>Of this small lute gave ease to Petrarch's wound. <hr:s  eID="s04"/>
  </l>
</lg>
```

Depending on how the modifications are carried out, this method may be TEI-conformant, may represent an extension of the TEI, or may produce a non-conformant document.

- The method is TEI-conformant if the modified elements and attributes can be mapped without loss of information to existing TEI markup structures such as milestone or anchor elements automatically (see *24.4. Conformance*).
- The method represents an Extension if the modified elements are placed in a distinct, non-TEI namespace (see *24.4. Conformance*).
- The method is non-conformant—and indeed strongly deprecated—if the modified elements and attributes are not placed in a distinct, non-TEI namespace (see *24.4.3. Conformance to the TEI Abstract Model*).

In each of the above examples (except the last), the relationship between the start and end delimiters (where these exist) of a given feature is implicit: it is assumed that ”end” delimiters close the nearest preceding ”start” delimiter, or, in the case of milestones, that the milestone marks both the end of the preceding example and the beginning of the next. Complications arise, however, when the non-nesting text overlaps with other non-nesting text of the same type, as, for example, in a grammatical analysis of the various possible interpretations of the noun phrase *fast trains and planes*. In this case, the adjective *fast* can be understood as either modifying *trains and planes* or just *trains*:

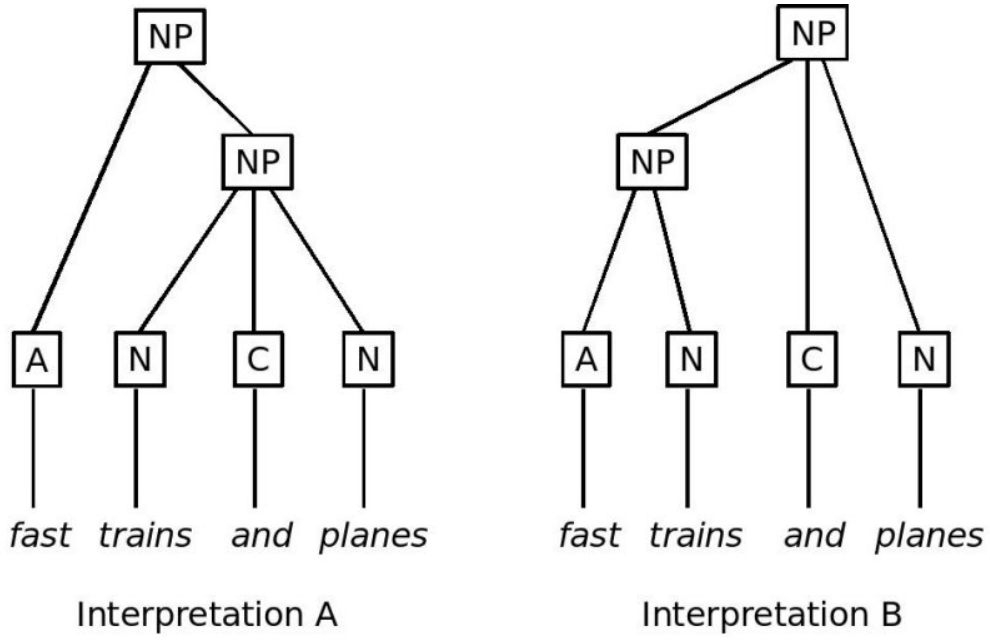



Figure 35: Two interpretations of the phrase *Fast trains and planes*

In order to encode the possible analyses of this phrase, an unambiguous method of associating opening and closing segment boundary delimiters is required:

```
<phr  function="NP">
  <anchor  type="delimiter"  subtype="NPstart"
    xml:id="NPInterpretationB"/>
  <w  function="A">Fast</w>
```

```
  <anchor  type="delimiter"  subtype="NPstart"
    xml:id="NPInterpretationA"/>
  <w  function="N">trains</w>
  <anchor  type="delimiter"  subtype="NPend"
    corresp="#NPInterpretationB"/>
  <w  function="C">and</w>
  <w  function="N">planes</w>
  <anchor  type="delimiter"  subtype="NPend"
    corresp="#NPInterpretationA"/>
</phr>
```

In this encoding, the first interpretation, in which *fast* modifies the NP *trains and planes*, the NP *trains and planes* is opened using an <anchor> tag with the *xml:id* value *NPInterpretationA* and closed with an <anchor> with the same value on *corresp*; in the second interpretation, in which *fast* forms a NP with *trains*, the NP *fast cars* is opened using an <anchor> tag with the *xml:id* value *NPInterpretationB* and closed with an <anchor> tag that has the same value on *corresp*.

Despite their advantages, segment boundary delimiters incur the disadvantage of cumbersome processing: since the elements of the analysis (e.g., the sentences in the poems, or phrases in the above example) are not uniformly represented by nodes in the document tree, they must be reconstituted by software in an ad hoc fashion, which is likely to be difficult and may be error prone.

Most important for some encoders, the method also disguises the relationship between the beginning and the ending of each logical element. This makes it impossible for standard validation software to provide the same kind of validation possible elsewhere in the encoding. When using grammar-based schema languages it is not possible to define a content model for the range limited by empty elements.[126]

## 21.3 Fragmentation and Reconstitution of Virtual Elements

A third method involves breaking what might be considered a single logical (but non-nesting) element into multiple smaller structural elements that fit within the dominant hierarchy but can be reconstituted virtually. For example, if a passage of direct discourse begins in the middle of one paragraph and continues for several more paragraphs, one could encode the passage as a series of <said> elements, each fitting within a <p> element. The resulting encoding is valid XML, but the text in each <said> element represents only a portion of the complete passage of direct discourse. For this reason these elements are sometimes called 'partial elements'.

In the case of our selection from Pinsky's poem, for example, the second passage of direct quotation, which crosses a line boundary and is broken up by a *She said* in the narrator's voice, can be made to fit within the hierarchy established by the metrical lineation by using two <said> elements:

```
<lg>
  <l>Catholic woman of twenty-seven with five children</l>
  <l>And a first-rate body—pointed her finger</l>
  <l>at the back of one certain man and asked me,</l>
  <l>
    <said  n="quotation1">Is that guy a psychiatrist?</said> and by god he was!
  <said  n="quotation2">Yes,</said>
  </l>
  <l>She said, <said  n="quotation2">He <emph>looks</emph> like a
        psychiatrist.</said>
  </l>
  <l>Grown quiet, I looked at his pink back, and thought.</l>
</lg>
```

[126] Grammar based schema languages (e.g., DTD, W3C Schema, and RELAX NG) are used to define markup languages (e.g., XHTML or TEI). Rule-based schema languages (e.g., Schematron) can be used to define further constraints. Such a rule-based schema language permits a sequence of certain elements between empty elements to be legitimized or prohibited.

Similarly, the sentences in our example from Wordsworth could be encoded:

```
<l>
  <seg  n="sentence1">Scorn not the sonnet;</seg>
  <seg  n="sentence2">critic, you have frowned,</seg>
</l>
<l>
  <seg  n="sentence2">Mindless of its just honours;</seg>
  <seg  n="sentence3">with this key</seg>
</l>
<l>
  <seg  n="sentence3">Shakespeare unlocked his heart;</seg>
  <seg  n="sentence4">the melody</seg>
</l>
<l>
  <seg  n="sentence4">Of this small lute gave ease to Petrarch's wound.</seg>
</l>
```

There are two main problems with this type of encoding. The first is that it invariably means that the encoding will have more elements claiming to represent a feature than there are actual instances of that feature in the text. Thus, for example, the passage from 'Scorn not the sonnet' marks seven spans of text using <seg>, even though there are only four linguistic sentences in the passage.

The second problem is that it can be semantically misleading. Although they are tagged using the element for *sentence*, for example, very few of the textual features encoded using <seg> in this example represent actual linguistic sentences: *with this key*, for example, is a prepositional phrase, not a sentence; *Of this small lute gave ease to Petrarch's wound* is a string corresponding to no single grammatical category.

Taken together, these problems can make automatic analysis of the fragmented features difficult. An analysis that intended to count the number of sentences in Wordsworth's poem, for example, would arrive at an inflated figure if it understood the <seg> elements to represent complete rhetorical sentences; if it wanted to do an analysis of his syntax, it would not be able to assume that <seg> delimited linguistic sentences.

The technique of fragmentation is often complemented by the technique of virtual joins. Virtual joins may be used to combine objects in the text to a new hierarchy. Here is 'Scorn not the sonnet' again; this time the relationship between the parts of the fragmented sentences is indicated explicitly using the *next* and *prev* attributes described in *17.7. Aggregation*.

```
<l>
  <seg>Scorn not the sonnet;</seg>
  <seg  next="#s2b"  xml:id="s2a">critic, you have frowned,</seg>
</l>
<l>
  <seg  prev="#s2a"  xml:id="s2b">Mindless of its just honours;</seg>
  <seg  next="#s3b"  xml:id="s3a">with this key</seg>
</l>
<l>
  <seg  prev="#s3a"  xml:id="s3b">Shakespeare unlocked his heart;</seg>
  <seg  next="#s4b"  xml:id="s4a">the melody</seg>
</l>
<l>
  <seg  prev="#s4a"  xml:id="s4b">Of this small lute gave ease to Petrarch's
wound.</seg>
</l>
```

This method of virtually joining partial elements is sometimes called 'chaining'.

For fragments encoded using <ab>, <l>, <lg>, , or elements that belong to the `att.segLike` class, an even simpler mechanism for virtually joining fragments exists: the use of the *part* attribute with the value *I* (Initial), *M*

(Medial), or *F* (Final) as described in *17.3. Blocks, Segments, and Anchors.* Here is the above example recoded to reflect this method:

```
<l>
  <seg>Scorn not the sonnet;</seg>
  <seg  part="I">critic, you have frowned,</seg>
</l>
<l>
  <seg  part="F">Mindless of its just honours;</seg>
  <seg  part="I">with this key</seg>
</l>
<l>
  <seg  part="F">Shakespeare unlocked his heart;</seg>
  <seg  part="I">the melody</seg>
</l>
<l>
  <seg  part="F">Of this small lute gave ease to Petrarch's wound.</seg>
</l>
```

This method is TEI-conformant and simple to use. Its disadvantage is that it does not work well for cases of self-overlap, or if there are nested occurrences of the same element type, as it can become difficult to ascertain which initial, medial, or final partial element should be combined with which others or in which order. This problem becomes evident if we attempt to combine a detailed Grammatical view of the Pinsky example with its metrical encoding:

```
<lg>
  <l>
    <seg  part="I">Catholic woman of twenty-seven with five children</seg>
  </l>
  <l>
    <seg  part="M">And a first-rate body—pointed her finger</seg>
  </l>
  <l>
    <seg  part="M">at the back of one certain man and asked me,</seg>
  </l>
  <l>
    <seg  part="F">"<seg>Is that guy a psychiatrist?</seg>" and by god he was!</seg>
    <seg  part="I">"<seg  part="I">Yes,</seg>"</seg>
  </l>
  <l>
    <seg  part="F">She said, "<seg  part="F">He <emph>looks</emph> like a
psychiatrist.</seg>"</seg>
  </l>
  <l>
    <seg>Grown quiet, I looked at his pink back, and thought.</seg>
  </l>
</lg>
```

A third method for aggregating fragmented partial elements involves using markup that is not directly part of the encoding, e.g., the <join> element. In this method, a <join> element is used elsewhere in the document to indicate explicitly the members of the virtual element:

```
<l>
  <w  xml:id="w01">Scorn</w>
```

```
  <w  xml:id="w02">not</w>
  <w  xml:id="w03">the</w>
  <w  xml:id="w04">sonnet</w>; <w  xml:id="w05">critic</w>, <w  xml:id="w06">you</w>
  <w  xml:id="w07">have</w>
  <w  xml:id="w08">frowned</w>,
</l>
<l>
  <w  xml:id="w09">Mindless</w>
  <w  xml:id="w10">of</w>
  <w  xml:id="w11">its</w>
  <w  xml:id="w12">just</w>
  <w  xml:id="w13">honours</w>; <w  xml:id="w14">with</w>
  <w  xml:id="w15">this</w>
  <w  xml:id="w16">key</w>
</l>
<l>
  <w  xml:id="w17">Shakespeare</w>
  <w  xml:id="w18">unlocked</w>
  <w  xml:id="w19">his</w>
  <w  xml:id="w20">heart</w>; <w  xml:id="w21">the</w>
  <w  xml:id="w22">melody</w>
</l>
<l>
  <w  xml:id="w23">Of</w>
  <w  xml:id="w24">this</w>
  <w  xml:id="w25">small</w>
  <w  xml:id="w26">lute</w>
  <w  xml:id="w27">gave</w>
  <w  xml:id="w28">ease</w>
  <w  xml:id="w29">to</w>
  <w  xml:id="w30">Petrarch's</w>
  <w  xml:id="w31">wound</w>.
</l>
<!-- Elsewhere in the document -->
<p>
  <join  result="s"  scope="root"
    target="#w01 #w02 #w03 #w04"/>
  <join  result="s"  scope="root"
    target="#w05 #w06 #w07 #w08 #w09 #w10 #w11 #w12 #w13"/>
  <join  result="s"  scope="root"
    target="#w14 #w15 #w16 #w17 #w18 #w19 #w20"/>
  <join  result="s"  scope="root"
    target="#w21 #w22 #w23 #w24 #w25 #w26 #w27 #w28 #w29 #w30 #w31"/>
</p>
```

This use of <join> is TEI-conformant.

The major advantage of fragmentation and virtual joins is that it allows all the hierarchies in the text to be handled explicitly: both the privileged one directly represented and the alternate hierarchy that has been split up and rejoined. The major disadvantages are that (like most of the other methods described here) it privileges one hierarchy over the others, requires special processing to reconstitute the elements of the other hierarchies, and, except in the case of <join>, can be semantically misleading.

## 21.4 Stand-off Markup

Most markup is characterized by the embedding of elements in the text. An alternative approach separates the text and the elements used to describe it. This approach is known as stand-off markup (see section *17.9. Stand-off Markup*). It establishes a new hierarchy by building a new tree whose nodes are XML elements that do not contain

textual content, but rather links to another *layer*: *a node in another XML document or a span of text.* This approach can be subdivided according to different criteria. A first distinction concerns the link base, i.e. the content to which annotations are to be applied. Sometimes the link target contains markup that can be referred to explicitly, as in the following example where the offset markup uses the *xml:id* values on <w> to provide targets for <xi:include>[127]:

```
<l>
  <w  xml:id="w001">Scorn</w>
  <w  xml:id="w002">not</w>
  <w  xml:id="w003">the</w>
  <w  xml:id="w004">sonnet</w>; <w  xml:id="w005">critic</w>, <w  xml:id="w006">you</w>
  <w  xml:id="w007">have</w>
  <w  xml:id="w008">frowned</w>,
</l>
<l>
  <w  xml:id="w009">Mindless</w>
  <w  xml:id="w010">of</w>
  <w  xml:id="w011">its</w>
  <w  xml:id="w012">just</w>
  <w  xml:id="w013">honours</w>; <w  xml:id="w014">with</w>
  <w  xml:id="w015">this</w>
  <w  xml:id="w016">key</w>
</l>
<l>
  <w  xml:id="w017">Shakespeare</w>
  <w  xml:id="w018">unlocked</w>
  <w  xml:id="w019">his</w>
  <w  xml:id="w020">heart</w>; <w  xml:id="w021">the</w>
  <w  xml:id="w022">melody</w>
</l>
<l>
  <w  xml:id="w023">Of</w>
  <w  xml:id="w024">this</w>
  <w  xml:id="w025">small</w>
  <w  xml:id="w026">lute</w>
  <w  xml:id="w027">gave</w>
  <w  xml:id="w028">ease</w>
  <w  xml:id="w029">to</w>
  <w  xml:id="w030">Petrarch's</w>
  <w  xml:id="w031">wound</w>.
</l>
<!-- elsewhere in the current document -->


<p xmlns:xi="http://www.w3.org/2001/XInclude">
<seg>
<xi:include xpointer="range(element(w001),element(w004))"/>
</seg>
<seg>
<xi:include xpointer="range(element(w005),element(w013))"/>
</seg>
<seg>
<xi:include xpointer="range(element(w014),element(w020))"/>
</seg>
<seg>
<xi:include xpointer="range(element(w021),element(w031))"/>
```

[127] A fake namespace is given for XInclude here, to avoid the markup being interpreted literally during processing.

```
</seg>
</p>
```

Note that the layer that uses XInclude to build another hierarchy might well be in another document, in which case the value of *href* of `<xi:xinclude>` would need to be the URL of the document that contains the base layer, in this case the <w> elements.

This is very similar to the use of <join> discussed above. The main advantages of the stand-off method are that it is possible to specify attributes on the aggregate <seg> elements, and that there exists off-the-shelf software that will perform appropriate processing. Stand-off markup may be used even when the base text being annotated is plain text, i.e. does not have any XML encoding. In this case, the range of text to be marked up is indicated by character offsets (see *17.2.4. TEI XPointer Schemes*, in particular *17.2.4.7. string-range()*). Another distinction concerns the number of files which can serve as link targets. Often, one (dedicated) annotation is used as the link target of all the other annotations. It is also possible to freely interlink several layers.

It has been noted that stand-off markup has several advantages over embedded annotations. In particular, it is possible to produce annotations of a text even when the source document is read-only. Furthermore, annotation files can be distributed without distributing the source text. Further advantages mentioned in the literature are that discontinuous segments of text can be combined in a single annotation, that independent parallel coders can produce independent annotations, and that different annotation files can contain different layers of information. Lastly, it has also been noted that this approach is elegant.

But there are also several drawbacks. First, new stand-off annotated layers require a separate interpretation, and the layers—although separate—depend on each other. Moreover, although all of the information of the multiple hierarchies is included, the information may be difficult to access using generic methods.

Inasmuch as it uses elements not included in the TEI namespace, stand-off markup involves an extension of the TEI.

## 21.5 Non-XML-based Approaches

There exist many non-XML methods of encoding a text that either solve or do not suffer the problem of the inability to encode overlapping hierarchies. These include, but are not limited to, the following proposals.

- Applying the notion of concurrent markup to XML (**?**). This reintroduces the CONCUR feature of SGML, which was omitted from the XML specification.
- Designing a form of document representation in which several trees share all or part of the same frontier, and in which each individual view of the document has the form of a tree (see **?**).
- The 'colored XML' proposal (**?**), which stores a body of information as a set of intertwined XML trees. This approach eliminates unnecessary redundancy and makes the database readily updatable, while allowing the user to exploit different hierarchical access paths.
- The MultiX proposal (**?**) , which represents documents as directed graphs. Because XML is used to represent the graph, the document is, at least in principle, manipulable with standard XML tools.
- The Just-In-Time-Trees proposal (**?**), which stores documents using XML, but processes the XML representation in non-standard ways and allows it to be mapped onto data structures that are different from those known from XML.
- The Layered Markup and Annotation Language ( LMNL) proposal. This offers alternatives to the basic XML linear form as well as its data and processing models. It uses an alternative notation to XML and a data structure based on Core Range Algebra (**?**).
- Markup Languages for Complex Documents ( MLCD). This provides a notation (TexMECS) and a data structure (Goddag) as well as a draft constraint language for the representation of non-hierarchical structures; see **?**.

These approaches are based either on non-standard XML processing or data models, or not based on XML at all. Since TEI is currently based on XML they are not described any further in these Guidelines. Use of these methods with the TEI will certainly involve extensions; in most cases the documents will also be non-conformant.

# 22 Certainty, Precision, and Responsibility

Encoders of text often find it useful to indicate that some aspects of the encoded text are problematic or uncertain, and to indicate who is responsible for various aspects of the markup of the electronic text. These Guidelines provide several methods of recording uncertainty about the text or its markup:

- the <note> element defined in section *3.9. Notes, Annotation, and Indexing* may be used with a value of certainty for its *type* attribute.
- the <certainty> element defined in this chapter may be used to record the nature and degree of the uncertainty in a more structured way.
- the <precision> element defined in this chapter may be used to record the accuracy with which some numerical value (such as a date or quantity) is provided by some other element or attribute.
- the <alt> element defined in the module for linking and segmentation may be used to provide alternative encodings for parts of a text, as described in section *17.8. Alternation*.

There are three methods of indicating responsibility for different aspects of the electronic text:

- the TEI header records who is responsible for an electronic text by means of the <respStmt> element and other more specific elements (<author>, <sponsor>, <funder>, <principal>, etc.) used within the <titleStmt>, <editionStmt>, and <revisionDesc> elements.
- the <note> element may be used with a value of resp or responsibility in its *type* attribute.
- the <respons> element defined in this chapter may be used to record fine-grained structured information about responsibility for individual tags in the text.

No special steps are needed to use the <note> and <respStmt> elements, since they are defined in the core module and header respectively. The <alt> element is only available when the module for linking has been selected, as described in chapter *17. Linking, Segmentation, and Alignment.* To use the <certainty>, <precision> or <respons> elements, the module for certainty and responsibility should be selected.

These three elements are all members of an attribute class called `att.scoping` from which they inherit the following attributes:

`att.scoping` provides attributes for selecting particular elements within a document.

- `@target` points at one or more sets of zero or more elements each.
- `@match` supplies an XPath selection pattern using the syntax defined in **?** which identifies a set of nodes, selected within the context identified by the *target* attribute if this is supplied, or within the context of the parent element if it is not.

These attributes enable statements about certainty, precision, or responsibility to be made with respect to the whole of a document, or any part or parts of it which can be identified using standard XML location methods. Several examples are given in the discussion of the <certainty> element below; the same mechanisms are available for all three elements discussed in this chapter.

## 22.1 Levels of Certainty

Many types of uncertainty may be distinguished. The <certainty> element is designed to encode the following sorts:

- a given tag may or may not correctly apply (e.g. a given word may be a personal name, or perhaps not)
- the precise point at which an element begins or ends is uncertain
- the value given for an attribute is uncertain
- the content given for an element is unreliable for any reason.

The following types of uncertainty are *not* indicated with the <certainty> element:

- the numerical precision associated with a number or date (for this use the <precision> element discussed in *22.2. Indications of Precision*)
- the content of the document being transcribed is identifiable, but may be read or understood in different ways (for this use the transcriptional elements such as <unclear>, discussed in chapter *12. Representation of Primary Sources*)

- a transcriber, editor, or author wishes to indicate a level of confidence in a factual assertion made in the text (for this use the interpretative mechanisms discussed in *18. Simple Analytic Mechanisms* and *19. Feature Structures*)

### 22.1.1 Using Notes to Record Uncertainty

The simplest way of recording uncertainty about markup is to attach a note to the element or location about which one is unsure. In the following (invented) paragraph, for example, an encoder might be uncertain whether to mark 'Essex' as a place name or a personal name, since both might be plausible in the given context:

> Elizabeth went to Essex. She had always liked Essex.

Using <note>, the uncertainty here may be recorded quite simply:

```
<persName>Elizabeth</persName> went to <placeName>Essex</placeName>. She had always
liked <placeName>Essex</placeName>.
<note  type="certainty"  resp="#MSM">It is not
clear here whether <mentioned>Essex</mentioned>
refers to the place or to the nobleman. -MSM</note>
```

Using the normal mechanisms, the note may be associated unambiguously with specific elements of the text, thus:

```
<persName>Elizabeth</persName> went to <placeName  xml:id="CE-p1a">Essex</placeName>.
She had always liked <placeName  xml:id="CE-p1b">Essex</placeName>.
<note  type="certainty"  resp="#MSM"
  target="#CE-p1a #CE-p1b">It
is not clear here whether <mentioned>Essex</mentioned>
refers to the place or to the nobleman. If the latter,
it should be tagged as a personal name. -<name  xml:id="MSM">Michael</name>
</note>
```

The advantage of this technique is its relative simplicity. Its disadvantage is that the nature and degree of uncertainty are not conveyed in any systematic way and thus are not susceptible to any sort of automatic processing.

### 22.1.2 Structured Indications of Uncertainty

To record uncertainty in a more structured way, susceptible of at least simple automatic processing, the <certainty> element may be used:

`<certainty>` indicates the degree of certainty associated with some aspect of the text markup.

- `@locus` indicates more exactly the aspect concerning which certainty is being expressed: specifically, whether the markup is correctly located, whether the correct element or attribute name has been used, or whether the content of the element or attribute is correct, etc.
- `@degree` indicates the degree of confidence assigned to the aspect of the markup named by the *locus* attribute.

Returning to the example, the <certainty> element may be used to record doubts about the proper encoding of 'Essex' in several ways of varying precision. To record merely that we are not certain that 'Essex' is in fact a place name, as it is tagged, we use the *target* attribute to identify the element in question, and the *locus* attribute to indicate which aspect of the markup we are uncertain about (in this case, whether we have used the correct 'name' for the element used to mark it):

```
Elizabeth went to
<placeName  xml:id="CE-pl1">Essex</placeName>.

<!-- ... elsewhere in the document ... -->
<certainty  target="#CE-pl1"  locus="name">
  <desc>possibly not a placename</desc>
</certainty>
```

There are no particular constraints as to where the <certainty> element is placed in a document; it may be placed adjacent to the target element, or elsewhere in the same or another document. Its position is however significant when the *target* attribute is not specified as further discussed below.

We may wish to record the probability, assessed in some subjective way, that 'Essex' really is a place name here. The *degree* attribute is used to indicate the degree of confidence associated with the <certainty> element, expressed as a number between 0 and 1:

```
<!-- ... --><certainty  target="#CE-pl1"  locus="name"
  degree="0.6"/>
```

This expresses the point of view that there is a 60 percent chance of 'Essex' being a place name here, and hence a 40 percent chance of its being a personal name. We can use two <certainty> elements to indicate the two probabilities independently. Both elements indicate the same location in the text, but the second provides an alternative choice of name identifier (in this case <persName>), which is given as the value of the *assertedValue* attribute:

```
<!-- ... --><certainty  target="#CE-pl1"  locus="name"
  degree="0.6">
  <desc>probably a placename, but possibly not</desc>
</certainty>
<certainty  target="#CE-pl1"  locus="name"
  degree="0.4"  assertedValue="persName">
  <desc>may refer to the Earl of Essex</desc>
</certainty>
```

In the simplest case, it is also possible to place the <certainty> element within the element concerned:

```
Elizabeth went to
<placeName>Essex
<certainty  locus="name"  degree="0.6"/>
</placeName>.
```

When no *target* is specified, by default the proposed certainty applies to its parent element, in this case the <placeName> element. The *match* attribute discussed below may be used to further vary this behaviour.

#### 22.1.2.1 Contingent Conditions

Finally, we may wish to make our probability estimates contingent on some condition. In the passage 'Elizabeth went to Essex; she had always liked Essex,' for example, we may feel there is a 60 percent chance that the county is meant, and a 40 percent chance that the earl is meant. But the two occurrences of the word are not independent: there is (we may feel) no chance at all that the first occurrence refers to the county and the second to the earl. We can express this by using the *given* attribute to list the identifiers of <certainty> elements.

```
Elizabeth went to <placeName  xml:id="CE-PL1">Essex</placeName>.
She had always liked <placeName  xml:id="CE-PL2">Essex</placeName>.

<!-- ... -->
<!-- 60% chance that P1 is a placename,
40% chance a personal name. -->
<certainty  xml:id="cert-1"  target="#CE-PL1"
  locus="name"  degree="0.6">
  <desc>probably a placename, but possibly not"</desc>
</certainty>
```

```
<certainty xml:id="cert-2" target="#CE-PL1"
  locus="name" assertedValue="persName" degree="0.4">
  <desc>may refer to the Earl of Essex"</desc>
</certainty>
<!-- 60% chance that P2 is a placename,
40% chance a personal name.
100% chance that it agrees with P1. -->
<certainty target="#CE-PL2" locus="name"
  given="#cert-1" degree="1.0">
  <desc>if CE-PL1 is a placename, CE-PL2 certainly is"</desc>
</certainty>
<certainty target="#CE-PL2" locus="name"
  assertedValue="persName" degree="1.0" given="#cert-2">
  <desc>if CE-PL1 is a personal name, then so is CE-PL2</desc>
</certainty>
```

When *given* conditions are listed, the <certainty> element is interpreted as claiming a given degree of confidence in a particular markup given the assertional content of the <certainty> elements indicated. That is, a conjectural assertion is being made solely on the assumption that the interpretation indicated by the element named by the *given* attribute is actually correct.

Conditional confidence may be less than 100 percent: given the sentence 'Ernest went to old Saybrook', we may interpret 'Saybrook' as a personal name or a place name, assigning a 60 percent probability to the former. If it is a place name, there may be a 50 percent chance that the place name actually in question is 'Old Saybrook' rather than 'Saybrook', while if it is correctly tagged as a personal name, it is much more likely (say, 90 percent certain) that the name is 'Saybrook'. Hence there is uncertainty about the correct location for the markup as well as about which markup to use. This state of affairs can be expressed using the <certainty> element thus:

```
Ernest went to <anchor xml:id="CE-a1"/> old <persName xml:id="CE-p2">Saybrook</persName>.

<certainty xml:id="cert1" target="#CE-p2"
  locus="name" degree="0.6"/>
<certainty target="#CE-p2" locus="start"
  given="#cert1" degree="0.9"/>
<certainty xml:id="cert2" target="#CE-p2"
  locus="name" assertedValue="placeName" degree="0.4"/>
<certainty target="#CE-p2" locus="start"
  given="#cert2" degree="0.5"/>
<certainty xml:id="cert3" target="#CE-p2"
  locus="start" assertedValue="#CE-a1" given="#cert1"
  degree="0.1"/>
<certainty xml:id="cert4" target="#CE-p2"
  locus="start" assertedValue="#CE-a1" given="#cert2"
  degree="0.5"/>
```

Note the use of the *assertedValue* on <certainty> elements cert3 and cert4 to reference the <anchor> element placed at the alternative starting point for the element.

Multiplying the numeric values out, this markup may be interpreted as assigning specific probabilities to three different ways of marking up the sentence:

```
Ernest went to old <persName>Saybrook</persName>. (0.6 * 0.9, or 0.54)
Ernest went to old <placeName>Saybrook</placeName>. (0.4 * 0.5, or 0.20)
Ernest went to <placeName>old Saybrook</placeName>. (0.4 * 0.5, or 0.20)
```

The probabilities do not add up to 1.00 because the markup indicates that if 'Saybrook' is (part of) a personal name, there is a 10 percent likelihood that the element should start somewhere other than the place indicated, without however giving an alternative location; there is thus a 6 percent chance (0.1 × 0.6) that none of the alternatives given is correct.

#### 22.1.2.2 Pervasive Conditions

We may also wish to indicate confidence in some aspect of the tagging throughout a document, rather than (as discussed so far) in one particular instance. The *match* attribute may be used to supply a pattern identifying the portion of a document concerning which certainty is being expressed. The value of the *match* attribute is an XSLT selection pattern using the syntax defined in **?**. In the following example, we wish to indicate a low degree of confidence that the <persName> elements used throughout the whole document have been correctly applied:

```
<certainty locus="name" degree="0.3"
  match="//persName"/>
```

No *target* has been supplied here, and so by default the <certainty> expressed would therefore apply to the parent element. However, in this case the XPath supplied as the value for *match* returns a set of all the <persName> elements in the document, independent of the current context. By contrast, in the following example

```
<div>
  <p>[...]</p>
</div>
<div>
  <certainty locus="name" degree="0.3"
    match=".//persName"/>
</div>
```

only the <persName> elements within the second  element are in question. Similarly, we may indicate that we have more confidence in the <persName> tagging within those  elements which have a *type* value of `checked`:

```
<certainty locus="name" degree="0.7"
  match="//div[@type='checked']//persName"/>
```

If an element in a document is matched by more than one match expression, then the most specific pattern applies. [128] As a simple case, if both the preceding <certainty> elements were present in the same document, a <persName> occurring within a  element would potentially match both pattern expressions. However because the second pattern is more specific than the former, in fact this is the only one that would apply. If multiple patterns match and have the same priority, then the first one (in document order) is applied. Only those statements of certainty which have matched in this sense are available for conditional application using the *given* attribute mentioned above.

When the *match* attribute is processed, the namespace bindings in force are those in effect at that point in the document. For example,

```
<div>
<!-- ... -->
  <certainty match=".//my:*" locus="value"
    degree="0.9"/>
</div>
```

might be used to indicate a high degree of certainty about the content of any elements taken the namespace associated with the prefix `my`. This namespace prefix must be associated with an appropriate namespace definition, either on the <certainty> element itself, or on one of its ancestor elements.

[128] Specificity of pattern matching is defined further in the XSLT3 reference cited above (see `https://www.w3.org/TR/xslt-30/#default-priority`)

#### 22.1.2.3 Content Uncertainty

Doubts about whether the content of an element is correct may also be expressed by assigning to *locus* the value *value*. For example, if the source is hard to read and so the transcription is uncertain:

```
I have a <emph  xml:id="CE-p3">bun</emph>.

<certainty  target="#CE-p3"  locus="value"
  degree="0.5"/>
```

Degrees of confidence in the proper expansion of abbreviations may also be expressed, as in the following example:

```
You will want to use
<choice>
  <expan  xml:id="CE-e1">Standard
     Generalized Markup Language</expan>
  <expan  xml:id="CE-e40">Some Grandiose Methodology for Losers</expan>
  <abbr>SGML</abbr>
</choice> ...

<!-- ... -->
<certainty  target="#CE-e1"  locus="value"
  degree="0.9"/>
<certainty  target="#CE-e40"  locus="value"
  degree="0.5"/>
```

The *assertedValue* attribute should be used to provide an alternative value for whatever aspect of the markup is in doubt: an alternative name, or the identifier of an alternative starting or ending point, as already shown, an alternative attribute value, or alternative element content, as in this example:

```
I have a <emph  xml:id="CE-P3">bun</emph>.

<certainty  target="#CE-P3"  locus="value"
  assertedValue="gun"  degree="0.8">
  <desc>a gun makes more sense in a holdup</desc>
</certainty>
```

Since attribute values have no internal substructure, the *assertedValue* attribute is not generally useful for specifying alternative transcriptions; it cannot for example be used if the alternative reading contains markup of any kind. More robust methods of handling uncertainties of transcription are the <unclear> element and the <app> and <rdg> elements described in chapter *13. Critical Apparatus*. The <certainty> element allows for indications of uncertainty to be structured with at least as much detail and clarity as appears to be currently required in most ongoing text projects.

#### 22.1.2.4 Target or Match?

As noted in *17. Linking, Segmentation, and Alignment*, the *target* attribute may take any general teidata.pointer as values and may thus also contain an XPath expression of arbitrary complexity. Because full support for XPath is not provided by current processors, it is not generally recommended TEI practice. There are however some simple cases in which XPath syntax is to be preferred, notably those in which the *xml:id* attribute is used to identify a single element occurrence. The usage #A (to indicate the element whose *xml:id* attribute has the value A) is syntactically much simpler than the equivalent xpath2 expression //*[@xml:id='A'] and is hence preferred throughout these guidelines.

For similar reasons, the <certainty> element may specify both a *target* value (expressed as an URI) and a *match* value (expressed as an XPath). The former defines the context within which the latter is to be evaluated. As previously

noted, if no value is supplied for *target*, the context within which the value of *match* should be evaluated is the parent element of the <certainty> element itself.

A typical case where it may be convenient to specify both *target* and *match* is that where we wish to indicate that the value of an attribute on some specific element is uncertain. In this case, the *locus* attribute takes the value value. For example, supposing there is only a 50 percent chance that the question was spoken by participant A:

```
<u xml:id="CE-u1" who="#A">Have you heard the election results?</u>
<certainty target="#CE-u1" match="@who"
 locus="value" degree="0.5"/>
```

or, equivalently and without the need to define a target,

```
<u who="#A">Have you heard the election results?<certainty match="@who" locus="value"
   degree="0.5"/>
</u>
```

The *match* and *target* attributes together provide a powerful mechanism which can be used to indicate precision for a large number of assertions throughout an encoded document in an economical way. Some further examples follow:

```
<certainty match="//p" locus="location"
 degree="0.2"/>
```

This encoding indicates that there is only a 0.2 certainty that the boundaries of all <p> elements in the document have been correctly identified.

```
<certainty target="#a101" match="p"
 locus="location" degree="0.2"/>
```

This encoding indicates that there is only a 0.2 certainty that the boundaries of the <p> elements contained by the element with *xml:id* value a101 have been correctly identified.

```
<persName resp="#LB">Essex
<certainty match="@resp" locus="value"
   degree="0.2"/>
</persName>
```

This encoding indicates that there is only a 0.2 certainty that the value for the *resp* attribute on the given <persName> element is correct.

```
<certainty match="//*/@resp" locus="value"
 degree="0.2"/>
```

This encoding indicates that there is only a 0.2 certainty that any value for the *resp* attribute is correct, wherever it appears in the document.

```
<certainty target="#dd001" match="@resp"
 locus="value" degree="0.2"/>
```

This encoding indicates that there is only a 0.2 certainty that the value for the *resp* attribute of the element indicated by the pointer #dd001 is correct

```
<certainty match="//*[@resp='#LB']"
  locus="value" degree="0.2"/>
```

This encoding indicates that there is only a 0.2 certainty that the content of any element the *resp* attribute of which has the value #LB is correct, wherever it appears in the document.

The <certainty> element and the other TEI mechanisms for indicating uncertainty provide a range of methods of graduated complexity. Simple expressions of uncertainty may be made by using the <note> element. This is simple and convenient, and can accommodate either a discursive and unstructured indication of uncertainty, or a complex and structured but probably project-specific expression of uncertainty. In general, however, unless special steps are taken, the <note> element does not provide as much expressive power as the <certainty> element, and in cases where highly structured certainty information are needed, it is recommended that the <certainty> element be preferred.

## 22.2 Indications of Precision

As noted above, certainty about the accuracy of an encoding or its content is not the same thing as the *precision* with which a value is specified. In the case of a date or a quantity, for example, we might be certain that the value given is imprecise, or uncertain about whether or not the value given is correct. The latter possibility would be represented by the <certainty> element discussed in the previous section; the former by the <precision> element discussed in this section.

The elements concerning which statements of precision are to be made are identified using the same *target* and *match* attributes inherited from the `att.scoping` class discussed in the previous section and in the same way. Other aspects are provided by other attributes as further discussed below.

`<precision>` indicates the numerical accuracy or precision associated with some aspect of the text markup.

- `@precision` characterizes the precision of the element or attribute pointed to by the <precision> element.
- `@stdDeviation` supplies a standard deviation associated with the value in question.

In *3.6.3. Numbers and Measures* several ways of indicating ranges of values were introduced. For example, if we know that a date falls between 1930 and 1935, without being certain exactly where, this fact may be encoded using attributes *notBefore* and *notAfter*, as in the following example:

```
<date notBefore="1930" notAfter="1935">Early in the 1930s</date>...
```

Equally, if we know that every page of a manuscript has a width of at least 10 cm but no more than 30, we can use the attributes *atLeast* and *atMost*, as in the following examples:

```
<width atLeast="10" atMost="30" unit="cm"
  scope="all"/>
```

Suppose however that the precision with which the value of such an attribute can be specified is variable. For example, suppose an event is dated 'about fifty years after the death of Augustus'. In this case, the precision of one end of the range (the death of Augustus) is higher than the other, assuming we know when Augustus died. We can say that the latest possible date is probably 50 years after that, but with less confidence than we can attach to the earliest possible date.

The <precision> element allows us to indicate the two attributes concerned and attach different levels of precision to them, using a similar mechanism as that provided for the <certainty> element:

```
<date xml:id="d001" notBefore="0014"
 notAfter="0064">About 50
years after the death of Augustus</date>
<precision target="#d001" match="@notAfter"
 precision="low"/>
<precision target="#d001"
 match="@notBefore" precision="high"/>
```

In much the same way, we may wish to indicate different levels of precision about the dating of either end of a historical period. For example, the elements defined for encoding personal data all bear a similar set of attributes to indicate normalized values for earliest or latest dates, etc. (see section *14.1.2. Dating Attributes*); the precision of these attribute values may be indicated in exactly the same way. For example,

```
<residence from="1857-03-01"
 notAfter="1857-04-30">From the 1st of March to
some time in April of 1857.
<precision match="@notAfter"
   precision="medium"/>
</residence>
```

It may also be useful to indicate that the precisions given for minimum and maximum quanta differ. For example, to indicate that all pages measure at least 10 cm wide, and at most *about* 30:

```
<width xml:id="w00t" atLeast="10"
 atMost="30" unit="cm" scope="all"/>
<precision target="#w00t" match="@atMost"
 precision="low"/>
```

The *stdDeviation* attribute may be used to indicate the standard deviation for a range of values. The generic <dim> element introduced in *11.3.4. Dimensions* might be used to record the average number of characters per line in a typescript. If in addition we wish to record the standard deviation for the values summarized by that average, this would require an additional <precision> element, as in the following example:

```
<dim xml:id="dim1" type="avgLineLength"
 unit="chars" quantity="62.4"/>
<precision target="#dim1" stdDeviation="4"/>
```

## 22.3 Attribution of Responsibility

In general, attribution of responsibility for the transcription and markup of an electronic text is made by <respStmt> elements within the header: specifically, within the title statement, the edition statement(s), and the revision history.

In some cases, however, more detailed element-by-element information may be desired. For example, an encoder may wish to distinguish between the individuals responsible for transcribing the content and those responsible for determining that a given word or phrase constitutes a proper noun. Where such fine-grained attribution of responsibility is required, the <respons> element can be used.

`<respons>` (responsibility) identifies the individual(s) responsible for some aspect of the content or markup of particular element(s).

`@locus` indicates the specific aspect of the encoding (markup or content) for which responsibility is being assigned.

This element allows one or more aspects of the markup to be attributed to a given individual. This element inherits the *target* and *match* attributes from the `att.scoping` class, in the same way as the <certainty> and

<precision> elements. Its *locus* attribute functions in the same way as that on the <certainty> element (see *22.1. Levels of Certainty*). It inherits the *resp* and *cert* attributes from the `att.global.responsibility` class.

For example, the following encoding indicates that RC is responsible for transcribing an illegible word, and that PMWR is responsible for identifying that word as a proper noun, i.e. deciding to mark it with the <persName> element at the location indicated:

```
Ernest went to old
<persName xml:id="CE-p5" rend="it">Saybrook</persName>.

<!-- ... -->
<respons target="#CE-p5" locus="value"
  resp="#RC"/>
<respons target="#CE-p5"
  locus="name location" resp="#PMWR"/>
<list type="encoders">
  <item xml:id="PMWR"/>
  <item xml:id="RC"/>
</list>
```

Similarly, in the following example, we indicate that RC is responsible for proposing the value of the *rend* attribute:

```
<respons target="#CE-p5" match="@rend"
  locus="value" resp="#RC"/>
```

## 22.4 The Certainty Module

The module described in this chapter makes available the following additional elements:

**Module certainty: Certainty, Precision, and Responsibility**

- *Elements defined*: certainty precision respons

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema*.

# 23 Documentation Elements

This chapter describes a module which may be used for the documentation of the XML elements and element classes which make up any markup scheme, in particular that described by the TEI Guidelines, and also for the automatic generation of schemas or DTDs conforming to that documentation. It should be used also by those wishing to customize or modify these Guidelines in a conformant manner, as further described in chapters *24.3. Customization* and *24.4. Conformance* and may also be useful in the documentation of any other comparable encoding scheme, even though it contains some aspects which are specific to the TEI and may not be generally applicable.

An overview of the kind of processing environment envisaged for the module described by this chapter may be helpful. In the remainder of this chapter we refer to software which provides such a processing environment as an *ODD processor*.[129] Like any other piece of XML software, an ODD processor may be instantiated in many ways: the current system uses a number of XSLT stylesheets which are freely available from the TEI, but this specification makes no particular assumptions about the tools which will be used to provide an ODD processing environment.

As the name suggests, an ODD processor uses a single XML document to generate multiple outputs. These outputs will include:

- formal reference documentation for elements, attributes, element classes, patterns, etc. such as those provided in *Appendix C Elements* below;
- detailed descriptive documentation, embedding some parts of the formal reference documentation, such as the tag description lists provided in this and other chapters of these Guidelines;
- declarative code for one or more XML schema languages, such as RELAX NG, W3C Schema, ISO Schematron, or DTD.

The input required to generate these outputs consists of running prose, and special purpose elements documenting the components (elements, classes, etc.) which are to be declared in the chosen schema language. All of this input is encoded in XML using elements defined in this chapter. In order to support more than one schema language, these elements constitute a comparatively high-level model which can then be mapped by an ODD processor to the specific constructs appropriate for the schema language in use. Although some modern schema languages such as RELAX NG or W3C Schema natively support self-documentary features of this kind, we have chosen to retain the ODD model, if only for reasons of compatibility with earlier versions of these Guidelines. For reasons of backwards compatibility, the ISO standard XML schema language RELAX NG (`http://www.relaxng.org`) may be used as a means of declaring content models and datatypes, but it is also possible to express content models using native TEI XML constructs. We also use the ISO Schematron language to define additional constraints beyond those expressed in the content model, as further discussed in *23.5.2. Additional Constraints* below.

In the TEI system, a *schema* is built by combining element and attribute declarations, more or less as required. Each element is documented by an appropriate *specification element* and has an identifier unique across the whole TEI scheme. For convenience, these specifications are grouped into a number of discrete *modules*, which can also be combined more or less as required. Each major chapter of these Guidelines defines a distinct module. Each module declares a number of *elements* specific to that module, and may also populate particular *classes*. All classes are available globally, irrespective of the module in which they are declared; particular modules extend the meaning of a class by adding elements or attributes to it. Wherever possible, element content models are defined in terms of classes rather than in terms of specific elements. Modules can also declare particular *patterns*, which act as short-cuts for commonly used content models or class references.

In the present chapter, we discuss the components needed to support this system. In addition, section *23.1. Phrase Level Documentary Elements* discusses some general purpose elements which may be useful in any kind of technical documentation, wherever there is need to talk about technical features of an XML encoding such as element names and attributes. Section *23.2. Modules and Schemas* discusses the elements which are used to document XML *modules* and their high-level components. Section *23.3. Specification Elements* discusses the elements which document XML elements and their attributes, element classes, and generic patterns or macros. Finally, section *23.9. Module for Documentation Elements* provides a summary overview of the elements provided by this module.

[129] ODD is short for 'One Document Does it all', and was the name invented by the original TEI Editors for the predecessor of the system currently used for this purpose. See further **?** and ?.

# 23.1 Phrase Level Documentary Elements

## 23.1.1 Phrase Level Terms

In any kind of technical documentation, the following phrase-level elements may be found useful for marking up strings of text which need to be distinguished from the running text because they come from some formal language:

`<code>` contains literal code from some formal language such as a programming language.

`@lang` (formal language) a name identifying the formal language in which the code is expressed.

`<ident>` (identifier) contains an identifier or name for an object of some kind in a formal language. <ident> is used for tokens such as variable names, class names, type names, function names etc. in formal programming languages.

Like other phrase-level elements used to indicate the semantics of a typographically distinct string, these are members of the `model.emphLike` class. They are available anywhere that running prose is permitted when the module defined by this chapter is included in a schema.

The <code> and <ident> elements are intended for use when citing brief passages in some formal language such as a programming language, as in the following example:

```
<p>If the variable <ident>z</ident> has a value of zero, a statement such as
<code>x=y/z</code> will
usually cause a fatal error.</p>
```

If the cited phrase is a mathematical or chemical formula, the more specific <formula> element defined by the `figures` module (*15.2. Formulæ and Mathematical Expressions*) may be more appropriate.

A further group of similar phrase-level elements is also defined for the special case of representing parts of an XML document:

`<att>` (attribute) contains the name of an attribute appearing within running text.

`<gi>` (element name) contains the name (generic identifier) of an element.

`<tag>` (tag) contains text of a complete start- or end-tag, possibly including attribute specifications, but excluding the opening and closing markup delimiter characters.

`<val>` (value) contains a single attribute value.

These elements constitute the `model.phrase.xml` class, which is also a subclass of `model.phrase`. They are also available anywhere that running prose is permitted when the module defined by this chapter is included in a schema.

As an example of the recommended use of these elements, we quote from an imaginary TEI working paper:

```
<p>The <gi>gi</gi> element is used to tag element
names when they appear in the text; the <gi>tag</gi> element however is used to show
how a tag as
such might appear. So one might talk of an occurrence of the <gi>blort</gi> element
which had been
tagged <tag>blort type='runcible'</tag>. The <att>type</att> attribute may take any
name token as
value; the default value is <val>spqr</val>, in memory of its creator.</p>
```

Within technical documentation, it is also often necessary to provide more extended examples of usage or to present passages of markup for discussion. The following special elements are provided for these purposes:

`<eg>` (example) contains any kind of illustrative example.

`<egXML>` (example of XML) a single XML fragment demonstrating the use of some XML, such as elements, attributes, or processing instructions, etc., in which the <egXML> element functions as the root element.

Like the <code> element, the <egXML> element is used to mark strings of formal code, or passages of XML markup. The <eg> element may be used to enclose any kind of example, which will typically be rendered as a distinct block, possibly using particular formatting conventions, when the document is processed. It is a specialized

form of the more general <q> element provided by the TEI core module. In documents containing examples of XML markup, the <egXML> element should be used for preference, as further discussed below in *23.4.2. Exemplification of Components*, since the content of this element can be checked for well-formedness.

These elements are added to the class `model.egLike` when this module is included in a schema. That class is a part of the general `model.inter` class, thus permitting <eg> or <egXML> elements to appear either within or between paragraph-like elements.

### 23.1.2 Element and Attribute Descriptions

Within the body of a document using this module, the following elements may be used to reference parts of the specification elements discussed in section *23.3. Specification Elements*, in particular the brief prose descriptions these provide for elements and attributes.

`<specList>` (specification list) marks where a list of descriptions is to be inserted into the prose documentation.

`<specDesc>` (specification description) indicates that a description of the specified element, class, or macro should be included at this point within a document.

`@atts` (attributes) supplies attribute names for which descriptions should additionally be obtained.

TEI practice recommends that a <specList> listing the elements under discussion introduce each subsection of a module's documentation. The source for the present section, for example, begins as follows:

```
<div>
  <head>Element and Attribute Descriptions</head>
  <p>Within the body of a document using this module, the following elements may be
used to reference parts of the specification elements
    discussed in section <ptr  target="#TDcrystals"/>, in particular the brief prose
descriptions these provide for elements and attributes.
  <specList>
     <specDesc  key="specList"/>
     <specDesc  key="specDesc"  atts="atts"/>
   </specList>
  </p>
  <p>TEI practice recommends that a <gi>specList</gi> listing the elements under … </p>
<!-- ... -->
</div>
```

When formatting the <ptr> element in this example, an ODD processor might simply generate the section number and title of the section referred to, perhaps additionally inserting a link to the section. In a similar way, when processing the <specDesc> elements, an ODD processor may recover relevant details of the elements being specified (<specList> and <specDesc> in this case) from their associated declaration elements: typically, the details recovered will include a brief description of the element and its attributes. These, and other data, will be stored in a specification element elsewhere within the current document, or they may be supplied by the ODD processor in some other way, for example from a database. For this reason, the link to the required specification element is always made using a TEI-defined key rather than an XML IDREF value. The ODD processor uses this key as a means of accessing the specification element required. There is no requirement that this be performed using the XML ID/IDREF mechanism, but there is an assumption that the identifier be unique.

A <specDesc> generates in the documentation the identifier, and also the contents of the <desc> child of whatever specification element is indicated by its *key* attribute, as in the example above. Documentation for any attributes specified by the *atts* attribute will also be generated as an associated attribute list.

## 23.2 Modules and Schemas

As mentioned above, the primary purpose of this module is to facilitate the documentation and creation of an XML schema derived from the TEI Guidelines. The following elements are provided for this purpose:

`<schemaSpec>` (schema specification) generates a TEI-conformant schema and documentation for it.

`<moduleSpec>` (module specification) documents the structure, content, and purpose of a single module, i.e. a named and externally visible group of declarations.

`<moduleRef>` (module reference) references a module which is to be incorporated into a schema.

`@include` supplies a list of the elements which are to be copied from the specified module into the schema being defined.

`@except` supplies a list of the elements which are not to be copied from the specified module into the schema being defined.

`<specGrp>` (specification group) contains any convenient grouping of specifications for use within the current module.

`<specGrpRef>` (reference to a specification group) indicates that the declarations contained by the <specGrp> referenced should be inserted at this point.

`<attRef>` (attribute pointer) points to the definition of an attribute or group of attributes.

`<elementRef>` points to the specification for some element which is to be included in a schema.

A *module* is a convenient way of grouping together element and other declarations, and of associating an externally-visible name with the resulting group. A *specification group* performs essentially the same function, but the resulting group is not accessible outside the scope of the ODD document in which it is defined, whereas a module can be accessed by name from any TEI schema specification. Elements, and their attributes, element classes, and patterns are all individually documented using further elements described in section *23.3. Specification Elements* below; part of that specification includes the name of the module to which the component belongs.

An ODD processor generating XML DTD or schema fragments from a document marked up according to the recommendations of this chapter will generate such fragments for each <moduleSpec> element found. For example, the chapter documenting the TEI module for names and dates contains a module specification like the following:

```
<moduleSpec  ident="namesdates">
  <idno  type="FPI">Names and Dates</idno>
  <desc>Additional elements for names and dates</desc>
</moduleSpec>
```

together with specifications for all the elements, classes, and patterns which make up that module, expressed using <elementSpec>, <classSpec>, or <macroSpec> elements as appropriate. (These elements are discussed in section *23.3. Specification Elements* below.) Each of those specifications carries a *module* attribute, the value of which is `namesdates`. An ODD processor encountering the <moduleSpec> element above can thus generate a schema fragment for the TEI `namesdates` module that includes declarations for all the elements (etc.) which reference it.

In most realistic applications, it will be desirable to combine more than one module together to form a complete *schema*. A schema consists of references to one or more modules or specification groups, and may also contain explicit declarations or redeclarations of elements (see further *23.8.1. TEI customizations*). Any combination of modules can be used to create a schema [130]

A schema can combine references to TEI modules with references to other (non-TEI) modules using different namespaces, for example to include mathematical markup expressed using MathML in a TEI document. By default, the effect of combining modules is to allow all of the components declared by the constituent modules to coexist (where this is syntactically possible: where it is not—for example, because of name clashes—a schema cannot be generated). It is also possible to over-ride declarations contained by a module, as further discussed in section *23.8.1. TEI customizations*

It is often convenient to describe and operate on sets of declarations smaller than the whole, and to document them in a specific order: such collections are called *specGrps* (specification groups). Individual <specGrp> elements are identified using the global *xml:id* attribute, and may then be referenced from any point in an ODD document using the <specGrpRef> element. This is useful if, for example, it is desired to describe particular groups of elements in a specific sequence. Note however that the order in which element declarations appear within the schema code generated from an ODD file element is not in general affected by the order of declarations within a <specGrp>.

An ODD processor will generate a piece of schema code corresponding with the declarations contained by a <specGrp> element in the documentation being output, and a cross-reference to such a piece of schema code when processing a <specGrpRef>. For example, if the input text reads

[130] The distinction between base and additional tagsets in earlier versions of the TEI scheme has not been carried forward into P5.

```
<p>This module contains three red elements: <specGrp  xml:id="RED">
   <elementSpec  ident="beetroot">
<!-- ... -->
   </elementSpec>
   <elementSpec  ident="east">
<!-- ... -->
   </elementSpec>
   <elementSpec  ident="rose">
<!-- ... -->
   </elementSpec>
  </specGrp> and two blue ones: <specGrp  xml:id="BLUE">
   <elementSpec  ident="sky">
<!-- ... -->
   </elementSpec>
   <elementSpec  ident="bayou">
<!-- ... -->
   </elementSpec>
  </specGrp>
</p>
```

then the output documentation will replace the two <specGrp> elements above with a representation of the schema code declaring the elements `<beetroot>`, `<east>`, and `<rose>` and that declaring the elements `<sky>` and `<bayou>` respectively. Similarly, if the input text contains elsewhere a passage such as

```
<div>
  <head>An overview of the imaginary module</head>
  <p>The imaginary module contains declarations for coloured things:
<specGrpRef  target="#RED"/>
   <specGrpRef  target="#BLUE"/>
  </p>
</div>
```

then the <specGrpRef> elements may be replaced by an appropriate piece of reference text such as 'The RED elements were declared in section 4.2 above', or even by a copy of the relevant declarations. As stated above, the order of declarations within the imaginary module described above will not be affected in any way. Indeed, it is possible that the imaginary module will contain declarations not present in any specification group, or that the specification groups will refer to elements that come from different modules. Specification groups are always local to the document in which they are defined, and cannot be referenced externally (unlike modules).

## 23.3 Specification Elements

The following elements are used to specify elements, classes, patterns, and datatypes:

`<elementSpec>` (element specification) documents the structure, content, and purpose of a single element type.

`<classSpec>` (class specification) contains reference information for a TEI element class; that is a group of elements which appear together in content models, or which share some common attribute, or both.

`@generate` indicates which alternation and sequence instantiations of a model class may be referenced. By default, all variations are permitted.

`<macroSpec>` (macro specification) documents the function and implementation of a pattern.

`<dataSpec>` (datatype specification) documents a datatype.

Unlike most elements in the TEI scheme, each of these 'specification elements' has a fairly rigid internal structure consisting of a large number of child elements which are always presented in the same order. Furthermore, since these elements all describe markup objects in broadly similar ways, they have several child elements in common. In

the remainder of this chapter, we discuss first the elements which are common to all the specification elements, and then those which are specific to a particular type.

Specification elements may appear at any point in an ODD document, both between and within paragraphs as well as inside a <specGrp> element, but the specification element for any particular component may only appear once (except in the case where a modification is being defined; see further *23.8.1. TEI customizations*). The order in which they appear will not affect the order in which they are presented within any schema module generated from the document. In documentation mode, however, an ODD processor will output the schema declarations corresponding with a specification element at the point in the text where they are encountered, provided that they are contained by a <specGrp> element, as discussed in the previous section. An ODD processor will also associate all declarations found with the nominated module, thus including them within the schema code generated for that module, and it will also generate a full reference description for the object concerned in a catalogue of markup objects. These latter two actions always occur irrespective of whether or not the declaration is included in a <specGrp>.

## 23.4 Common Elements

This section discusses the child elements common to all of the specification elements; some of these are defined in the core module (*3.4.1. Terms and Glosses*). These child elements are used to specify the naming, description, exemplification, and classification of the specification elements.

### 23.4.1 Description of Components

`<gloss>` (gloss) identifies a phrase or word used to provide a gloss or definition for some other word or phrase.

`<desc>` (description) contains a short description of the purpose, function, or use of its parent element, or when the parent is a documentation element, describes or defines the object being documented.

`<equiv>` (equivalent) specifies a component which is considered equivalent to the parent element, either by co-reference, or by external link.

- `@uri` (uniform resource identifier) references the underlying concept of which the parent is a representation by means of some external identifier.
- `@filter` references an external script which contains a method to transform instances of this element to canonical TEI.
- `@name` a single word which follows the rules defining a legal XML name (see `https://www.w3.org/TR/REC-xml/#dt-name`), naming the underlying concept of which the parent is a representation.
- `@predicate [att.predicate]` the condition under which the element bearing this attribute applies, given as an XPath predicate expression.

`<altIdent>` (alternate identifier) supplies the recommended XML name for an element, class, attribute, etc. in some language.

`<listRef>` (list of references) supplies a list of significant references in the current document or elsewhere.

`<remarks>` (remarks) contains any commentary or discussion about the usage of an element, attribute, class, or entity not otherwise documented within the containing element.

The <gloss> element may be used to provide a brief explanation for the name of the object if this is not self-explanatory. For example, the specification for the element <ab> used to mark arbitrary blocks of text begins as follows:

```
<elementSpec module="linking" ident="ab">
  <gloss>anonymous block</gloss>
<!--... -->
</elementSpec>
```

A <gloss> may also be supplied for an attribute name or an attribute value in similar circumstances:

```
<valList  type="open">
  <valItem  ident="susp">
    <gloss>suspension</gloss>
    <desc>the abbreviation provides the first letter(s) of the word or phrase, omitting
the
        remainder.</desc>
  </valItem>
  <valItem  ident="contr">
    <gloss>contraction</gloss>
    <desc>the abbreviation omits some letter(s) in the middle.</desc>
  </valItem>
<!--...-->
</valList>
```

Note that the <gloss> element is needed to explain the significance of the identifier for an item only when this is not apparent, for example because it is abbreviated, as in the above example. It should not be used to provide a full description of the intended meaning (this is the function of the <desc> element), nor to comment on equivalent values in other schemes (this is the purpose of the <equiv> element), nor to provide alternative versions of the *ident* attribute value in other languages (this is the purpose of the <altIdent> element).

The contents of the <desc> element provide a brief characterization of the intended function of the object being documented in a form that permits its quotation out of context, as in the following example:

```
<elementSpec  module="core"  ident="foreign">
<!--... -->
  <desc  xml:lang="en"
    versionDate="2007-07-21">identifies a word or phrase as belonging to some
     language other than that of the surrounding text. </desc>
<!--... -->
</elementSpec>
```

By convention, a <desc> element begins with a verb such as *contains*, *indicates*, *specifies*, etc. and contains a single clause.

Both the <gloss> and <desc> elements (in addition to <exemplum>, <remarks>, and <valDesc>) are members of `att.translatable`, and thus carry the *versionDate* attributre. Where specifications are supplied in multiple languages, these elements may be repeated as often as needed. Each such element should carry both an *xml:lang* and a *versionDate* attribute to indicate the language used and the date on which the translated text was last checked against its source.

The <equiv> element is used to document equivalencies between the concept represented by this object and the same concept as described in other schemes or ontologies. The *uri* attribute is used to supply a pointer to some location where such external concepts are defined. For example, to indicate that the TEI <death> element corresponds to the concept defined by the CIDOC CRM category E69, the declaration for the former might begin as follows:

```
<elementSpec  module="namesdates"
  ident="death">
  <equiv  name="E69"
    uri="http://cidoc.ics.forth.gr/"/>
  <desc>
<!--... -->
  </desc>
</elementSpec>
```

The <equiv> element may also be used to map newly-defined elements onto existing constructs in the TEI, using the *filter* and *name* attributes to point to an implementation of the mapping. This is useful when a TEI customization

(see *24.3. Customization*) defines 'shortcuts' for convenience of data entry or markup readability. For example, suppose that in some TEI customization an element <bo> has been defined which is conceptually equivalent to the standard markup construct <hi rend='bold'>. The following declarations would additionally indicate that instances of the <bo> element can be converted to canonical TEI by obtaining a filter from the URI specified, and running the procedure with the name bold. The *mimeType* attribute specifies the language (in this case XSL) in which the filter is written:

```
<elementSpec ident="bo"
  ns="http://www.example.com/ns/nonTEI">
  <equiv filter="http://www.example.com/equiv-filter.xsl"
    mimeType="text/xsl" name="bold"/>
  <gloss>bold</gloss>
  <desc>contains a sequence of characters rendered in a bold face.</desc>
<!-- ... -->
</elementSpec>
```

The <altIdent> element is used to provide an alternative name for an object, for example using a different natural language. Thus, the following might be used to indicate that the <abbr> element should be identified using the German word *Abkürzung*:

```
<elementSpec ident="abbr" mode="change">
  <altIdent xml:lang="de">Abkürzung</altIdent>
<!--...-->
</elementSpec>
```

In the same way, the following specification for the <graphic> element indicates that the attribute *url* may also be referred to using the alternate identifier href:

```
<elementSpec ident="graphic" mode="change">
  <attList>
    <attDef mode="change" ident="url">
      <altIdent>href</altIdent>
    </attDef>
<!-- ... -->
  </attList>
</elementSpec>
```

By default, the <altIdent> of a component is identical to the value of its *ident* attribute.

The <remarks> element contains any additional commentary about how the item concerned may be used, details of implementation-related issues, suggestions for other ways of treating related information etc., as in the following example:

```
<elementSpec module="core" ident="foreign">
<!--... -->
  <remarks>
    <p>This element is intended for use only where no other element is available to
mark the phrase
        or words concerned. The global <att>xml:lang</att> attribute should be used in
preference to
        this element where it is intended to mark the language of the whole of some
text element.</p>
    <p>The <gi>distinct</gi> element may be used to identify phrases belonging to
```

```
sublanguages or
        registers not generally regarded as true languages.</p>
  </remarks>
<!--... -->
</elementSpec>
```

A specification element will usually conclude with a list of references, each tagged using the standard <ptr> element, and grouped together into a <listRef> element: in the case of the <foreign> element discussed above, the list is as follows:

```
<listRef>
  <ptr  target="#COHQHF"/>
</listRef>
```

where the value COHQF is the identifier of the section in these Guidelines where this element is fully documented.

### 23.4.2 Exemplification of Components

`<exemplum>` (exemplum) groups an example demonstrating the use of an element along with optional paragraphs of commentary.

`<eg>` (example) contains any kind of illustrative example.

`<egXML>` (example of XML) a single XML fragment demonstrating the use of some XML, such as elements, attributes, or processing instructions, etc., in which the <egXML> element functions as the root element.

`@valid` indicates the intended validity of the example with respect to a schema.

`@source [att.global.source]` specifies the source from which some aspect of this element is drawn.

The <exemplum> element is used to combine a single illustrative example with an optional paragraph of commentary following or preceding it. The illustrative example itself may be marked up using either the <eg> or the <egXML> element.

The *source* attribute may be used on either element to indicate the source from which an example is taken, typically by means of a pointer to an entry in an associated bibliography, as in the following example:

```
  <exemplum versionDate="2008-04-06" xml:lang="fr">
  <p>L'element <gi>foreign</gi> s'applique également aux termes considerés étrangers.</p>
    <egXML xmlns="http://www.tei-c.org/ns/Examples" source="#fr-ex-Queneau_Journ">
      <p>Pendant ce temps-là, dans le bureau du rez- de-chaussée, les secrétaires faisaient du
        <foreign xml:lang="en">hulla-hoop</foreign>.</p>
    </egXML>
  </exemplum>
```

When, as here, an example contains valid XML markup, the <egXML> element should be used. In such a case, it will clearly be necessary to distinguish the markup within the example from the markup of the document itself. In an XML environment, this is easily done by using a different name space for the content of the <egXML> element. For example:

```
  <p>The <gi>term</gi> element may be used
to mark any technical term, thus:
<egXML xmlns="http://www.tei-c.org/ns/Examples">
This <term>recursion</term> is
giving me a headache.</egXML></p>
```

Alternatively, the XML tagging within an example may be 'escaped', either by using entity references to represent the opening angle bracket, or by wrapping the whole example in a CDATA marked section:

```
  <p>The <gi>term</gi> element may be used
to mark any technical term, thus:
<egXML xmlns="http://www.tei-c.org/ns/Examples">
This &lt;term&gt;recursion&lt;/term&gt; is
giving me a headache.</egXML></p>
```

or, equivalently:

```
  <p>The <gi>term</gi> element may be used
to mark any technical term, thus:
<egXML xmlns="http://www.tei-c.org/ns/Examples"><![CDATA[
This <term>recursion</term> is
giving me a headache.]]></egXML></p>
```

However, escaping the markup in this way will make it impossible to validate, and should therefore generally be avoided.

If the XML contained in an example is not well-formed then it must either be enclosed in a CDATA marked section, or 'escaped' as above: this applies whether the <eg> or <egXML> is used. The *valid* attribute on <egXML> may be used to indicate the XML validity of the example with respect to some schema, as being valid, invalid, or feasibly valid.

The <egXML> element should not be used to tag non-XML examples: the general purpose <eg> or <q> elements should be used for such purposes.

### 23.4.3 Classification of Components

In the TEI scheme elements are assigned to one or more *classes*, which may themselves have subclasses. The following elements are used to indicate class membership:

`<classes>` (classes) specifies all the classes of which the documented element or class is a member or subclass.

`<memberOf>` specifies class membership of the documented element or class.

`@key` specifies the identifier for a class of which the documented element or class is a member or subclass.

The <classes> element appears within either the <elementSpec> or <classSpec> element. It specifies the classes of which the element or class concerned is a member by means of one or more <memberOf> child elements. Each such element references a class by means of its *key* attribute. Classes themselves are defined by the <classSpec> element described in section *23.6. Class Specifications* below.

For example, to show that the element <gi> is a member of the class `model.phrase.xml`, the <elementSpec> which documents this element contains the following <classes> element:

```
<classes>
  <memberOf  key="model.phrase.xml"/>
</classes>
```

## 23.5 Element Specifications

The <elementSpec> element is used to document an element type, together with its associated attributes. In addition to the elements listed above, it may contain the following subcomponents:

`<content>` (content model) contains a declaration of the intended content model for the element (or other construct) being specified.
  `@autoPrefix` controls whether or not pattern names generated in the corresponding RELAX NG schema source are automatically prefixed to avoid potential nameclashes.

`<constraintSpec>` (constraint on schema) contains a formal constraint, typically expressed in a rule-based schema language, to which a construct must conform in order to be considered valid
  `@scheme` supplies the name of the language in which the constraints are defined

`<attList>` (attribute list) contains documentation for all the attributes associated with this element, as a series of <attDef> elements.
  `@org` (organization) specifies whether only one (choice) or all (group) of the attributes in the list are available.

`<model>` describes the processing intended for a specified element.
  `@behaviour` names the process or function which this processing model uses in order to produce output.

These subcomponents are discussed in the following sections.

### 23.5.1 Defining Content Models

As described in *v.3.2 Content Models: an Example* and *Content Model*, the *content* of the element being defined — that is, what elements are allowed inside it, and in what order they are permitted — is described by its *content model*. The content model is defined by the <content> child of <elementSpec>. There are three distinctly different ways of specifying a content model:

- The content model can be described using TEI elements defined by this chapter, as discussed in *23.5.1.1. Defining Content Models: TEI* immediately below. Two such TEI elements that may be used to define a content model are <dataRef> and <valList>. But because these are most often used to define attribute values, they are discussed separately near the beginning and towards the end of *23.5.3.2. Value Specification*, respectively.
- Alternatively, and primarily for backwards compatibility, the content model may be expressed using a RELAX NG pattern. This is discussed in *23.5.1.2. Defining Content Models: RELAX NG*, below.
- Lastly, content models may be expressed using a schema language other than TEI or RELAX NG, but no further recommendations on doing so are provided by these Guidelines.

#### 23.5.1.1 Defining Content Models: TEI

In the simplest case, the content model of an element may be expressed using a single <empty> element as the only child of <content>. This describes the element being defined as *empty*, meaning a valid instance of said element can not have any content.[131]

`<empty>` indicates the presence of an empty node within a content model.

More commonly, one or more of the following elements are used to define a content model:

`<elementRef>` points to the specification for some element which is to be included in a schema.

`<anyElement>` indicates the presence of any elements in a content model.

`<classRef>` points to the specification for an attribute or model class which is to be included in a schema.

`<macroRef>` points to the specification for some pattern which is to be included in a schema.

An <elementRef> provides the name of an element which may appear at a certain point in a content model. An <anyElement> also asserts that an element may appear at a certain point in a content model, but rather than providing the name of a particular element type that may appear, any element regardless of its name may appear (and may have any attributes). A <classRef> provides the name of a model class, members of which may appear at a certain point in content model.[132] A <macroRef> provides the name of a predefined macro, the expansion of which is to be inserted at a certain point in a content model.

These three elements are all members of an attribute class which provides attributes that further modify their significance as follows:

`att.repeatable` provides attributes for the elements which define component parts of a content model.

[131] It would still be allowed to contain comments or processing instructions, as these are not considered part of the content model.

[132] The <classRef> element may be used to refer to attribute classes, but this should not be done within a <content>.

@*minOccurs* (minimum number of occurrences) indicates the smallest number of times this component may occur.

@*maxOccurs* (maximum number of occurrences) indicates the largest number of times this component may occur.

Additionally, two wrapper elements are provided to indicate whether the components listed as their children form a sequence or an alternation:

`<sequence>` indicates that the constructs referenced by its children form a sequence.

`<alternate>` indicates that the constructs referenced by its children form an alternation.

These two wrapper elements are also members of `att.repeatable`. References listed as children of <sequence> must appear in the order and cardinality specified. Only one of the references listed as children of <alternate> may appear, although the cardinality of the <alternate> itself applies. Thus the following fanciful content model permits either any number of <ptr> elements (except zero) or any number of <ref> elements (except zero); at least one element must be present, but having both a <ptr> and a <ref> would be invalid.

```
<content>
  <alternate>
    <elementRef  key="ptr"  minOccurs="1"
      maxOccurs="unbounded"/>
    <elementRef  key="ref"  minOccurs="1"
      maxOccurs="unbounded"/>
  </alternate>
</content>
```

However, the following content model permits any number of either <ptr> or <ref> elements (except zero); one element must be present, and having both <ptr> elements and <ref> elements (even intermixed) would be valid.

```
<content>
  <alternate  minOccurs="1"
    maxOccurs="unbounded">
    <elementRef  key="ptr"/>
    <elementRef  key="ref"/>
  </alternate>
</content>
```

The <sequence> and <alternate> elements may be used in combination with great expressive power. For example, in the following example, which might be imagined as a clean replacement for the content of the <choice> element, one and only one of the element pairs <sic> and <corr>, <orig> and <reg>, or <abbr> and <expan> is allowed.

```
<content>
  <alternate>
    <sequence>
      <elementRef  key="sic"/>
      <elementRef  key="corr"/>
    </sequence>
    <sequence>
      <elementRef  key="orig"/>
      <elementRef  key="reg"/>
    </sequence>
    <sequence>
      <elementRef  key="abbr"/>
      <elementRef  key="expan"/>
    </sequence>
  </alternate>
</content>
```

In the following example, which might be imagined as a clean replacement for the content of the <address> element, the encoder is given a choice of either:

- a single <street> followed by a single <placeName> followed by a single <postCode> followed by an optional <country>, or
- 2, 3, or 4 <addrLine> elements.

```
<content>
  <alternate>
    <sequence>
      <elementRef  key="street"/>
      <elementRef  key="placeName"/>
      <elementRef  key="postCode"/>
      <elementRef  key="country"  minOccurs="0"
        maxOccurs="1"/>
    </sequence>
    <elementRef  key="addrLine"  minOccurs="2"
      maxOccurs="4"/>
  </alternate>
</content>
```

In addition to expressing where certain elements, members of a class of elements, or constructs matching a predefined macro may occur inside an element, a content model may permit a string of zero or more Unicode characters to occur at a certain point in the content model. This is indicated by supplying the element <textNode> within the <content> element.

`<textNode>` indicates the presence of a text node in a content model.

If nothing but a <textNode> element is present inside a <content> element, valid instances of the element being defined may contain a sequence of zero or more Unicode characters, but may not contain any elements.[133]

#### 23.5.1.2 Defining Content Models: RELAX NG

Element content models may also be defined using RELAX NG patterns. Here is a very simple example

```
<content>
  <rng:text/>
</content>
```

The element within whose specification element this <content> element appears will have a content model which is expressed in RELAX NG as `text`, using the RELAX NG namespace. This model will be copied unchanged to the output when RELAX NG schemas are being generated. When an XML DTD is being generated, an equivalent declaration (in this case `(#PCDATA)`) will be output.

Here is a more complex example:

```
<content>
  <rng:group>
    <rng:ref  name="fileDesc"/>
    <rng:zeroOrMore>
      <rng:ref  name="model.teiHeaderPart"/>
    </rng:zeroOrMore>
    <rng:optional>
      <rng:ref  name="revisionDesc"/>
```

[133] This content model is not used very often in the TEI scheme. Because only Unicode characters are permitted, there is no way to record characters that are not (yet) represented in Unicode. Thus in TEI instead of <textNode> we often use a reference to `macro.xtext` which permits both Unicode characters and the <g> element.

```
    </rng:optional>
  </rng:group>
</content>
```

This is the content model for the <teiHeader> element, expressed in the RELAX NG syntax, which again is copied unchanged to the output during schema generation. The equivalent DTD notation generated from this is `(fileDesc, (%model.teiHeaderPart;)*, revisionDesc?)`.

The RELAX NG language does not formally distinguish element names, attribute names, class names, or macro names: all names are patterns which are handled in the same way, as the above example shows. Within the TEI scheme, however, different naming conventions are used to distinguish amongst the objects being named. Unqualified names (`fileDesc`, `revisionDesc`) are always element names. Names prefixed with `model.` or `att.` (e.g. `model.teiHeaderPart` and `att.typed`) are always class names. In DTD language, classes are represented by parameter entities (`%model.teiHeaderPart;` in the above example); see further *1. The TEI Infrastructure*.

The RELAX NG pattern names generated by an ODD processor by default include a special prefix, the default value for which is set using the *prefix* attribute on <schemaSpec>. The purpose of this is to ensure that the pattern name generated is uniquely identified as belonging to a particular schema, and thus avoid name clashes. For example, in a RELAX NG schema combining the TEI element <ident> with another element called <ident> from some other vocabulary, the former will be defined by a pattern called `TEI_ident` rather than simply `ident`. Most of the time, this behaviour is entirely transparent to the user; the one occasion when it is not will be where a content model (expressed using RELAX NG syntax) needs explicitly to reference either the TEI <ident> or the other one. In such a situation, the *autoPrefix* attribute on <content> may be used. For example, suppose that we wish to define a content model for <term> which permits either a TEI <ident> or the <ident> defined by some other vocabulary. A suitable content model would be generated from the following <content> element:

```
<content autoPrefix="false">
  <rng:choice>
    <rng:ref name="TEI_ident"/>
    <rng:ref name="ident"/>
  </rng:choice>
</content>
```

### 23.5.2 Additional Constraints

In addition to the <content> element, a set of general <constraintSpec> elements can be used to express rules about the validity of an element. Like some other specification elements, they are identifiable (using the *ident* attribute) in order that a TEI customization may override, delete, or change them individually. Each <constraintSpec> can be expressed in any notation which is found useful; the notation used must be recorded using the *scheme* attribute.

Schematron is an ISO standard (ISO/IEC 19757-3:2006) that defines a simple XML vocabulary for an 'assertion language' which provides a powerful way of expressing constraints on the content of any XML document in addition to those provided by other schema languages. Such constraints can be embedded within a TEI schema specification — including a TEI customization specification — using the methods exemplified in this chapter.[134] An ODD processor will typically process any <constraintSpec> elements in a TEI specification whose *scheme* attribute indicates that they are expressed in Schematron to create an ISO Schematron schema which may be used to validate document instances. The ISO Schematron schema may be a free-standing document, or may be embedded in the RELAX NG schema output of the ODD processor.

Constraints are generally used to model local rules which may be outside the scope of the target schema language. For example, in earlier versions of these Guidelines several constraints on the usage of the attributes of the TEI element <relation> were expressed informally as follows: 'only one of the attributes *active* and *mutual* may be

[134] Indeed, these *Guidelines* themselves include many additional constraints which are expressed in the ISO Schematron language using this mechanism. A conformant TEI document should respect these constraints, although automatic validation of them may not be possible for all processors.

supplied; the attribute *passive* may be supplied only if the attribute *active* is supplied.'. In the current version of the Guidelines, constraint specifications expressed as Schematron rules have been added, as follows:

```
<constraintSpec ident="ref-or-key-or-name"
  scheme="schematron"
   xmlns:sch="http://purl.oclc.org/dsdl/schematron">
  <constraint>
    <sch:rule context="tei:relation">
      <sch:assert test="@ref or @key or @name">One of the attributes @name, @ref, or
@key must be supplied</sch:assert>
    </sch:rule>
  </constraint>
</constraintSpec>
<constraintSpec ident="active-mutual"
  scheme="schematron">
  <constraint>
    <sch:rule context="tei:relation">
      <sch:report test="@active and @mutual">Only one of the attributes @active and
@mutual may be supplied</sch:report>
    </sch:rule>
  </constraint>
</constraintSpec>
<constraintSpec ident="active-passive"
  scheme="schematron">
  <constraint>
    <sch:rule context="tei:relation">
      <sch:report test="@passive and not(@active)">the attribute @passive may be
supplied only if the attribute @active is supplied</sch:report>
    </sch:rule>
  </constraint>
</constraintSpec>
```

Note that the 'tei:' prefix needs to be bound to the TEI namespace using the Schematron language `<ns>` element. This can be done using the <constraintDecl> element, which may appear in either the <encodingDesc> or <schemaSpec> element. A <constraintDecl> contains declarations specific to all the <constraintSpec> elements in the current TEI document whose *scheme* matches that of the <constraintDecl>.

```
<constraintDecl scheme="schematron"
  queryBinding="xslt3"
   xmlns:sch="http://purl.oclc.org/dsdl/schematron">
  <sch:ns prefix="tei"
    uri="http://www.tei-c.org/ns/1.0"/>
  <sch:ns prefix="xi"
    uri="http://www.w3.org/2001/XInclude"/>
  <sch:let name="edition"
    value="substring-before( /*/tei:teiHeader//tei:editionStmt/tei:edition/@n, '.')"/>
  <sch:let name="uses_old_encoding"
    value="$edition cast as xs:integer lt 3"/>
</constraintDecl>
```

In this example the Schematron query language binding is set (to xslt3) for use by a Schematron processor, two namespace prefixes are bound to namespace URIs for later use within <constraint> elements, and the `$uses_old_encoding` variable is set (to either `true()` or `false()`) so that Schematron assertions elsewhere in the TEI document can easily test whether the edition number of the document being checked is 3 or more. (Presumably in this project there are minor encoding differences between the older and newer editions which do not merit an entirely different schema, but do merit the occasional different constraint.)

The following two examples are written presuming that the <constraintDecl> above is in force, that is that the Schematron processor will use an xslt3 binding, that 'tei:' is bound to the TEI namespace, that 'xi:' is bound to the XInclude namespace, and that the variable `$uses_old_encoding` is defined as a boolean. The first example models the constraint that a TEI  must contain either no subdivisions or at least two of them, with the added complication that larger subdivisions are provided as separate sect*NN*.xml files read in using XInclude:

```
<constraintSpec ident="subclauses"
  scheme="schematron"
   xmlns:sch="http://purl.oclc.org/dsdl/schematron">
  <constraint>
    <sch:rule context="tei:div">
      <sch:report test="count( tei:div | xi:include[ contains( @href, 'sect') ] ) eq
1">if it contains any
           subdivisions, a division must contain at least two of them</sch:report>
    </sch:rule>
  </constraint>
</constraintSpec>
```

The second example demonstrates that Schematron rules are also useful where an application needs to enforce rules on attribute values, in this case checking that various types of <title> are provided in the TEI header:

```
<constraintSpec ident="introtitle"
  scheme="schematron"
   xmlns:sch="http://purl.oclc.org/dsdl/schematron">
  <constraint>
    <sch:rule context="tei:teiHeader">
      <sch:assert test="tei:fileDesc/tei:titleStmt/tei:title[@type='introductory']"> an
introductory component of the title is expected</sch:assert>
    </sch:rule>
  </constraint>
</constraintSpec>
<constraintSpec ident="maintitle"
  scheme="schematron">
  <constraint>
    <sch:rule context="tei:teiHeader[ $uses_old_encoding ]">
      <sch:assert test="tei:fileDesc/tei:titleStmt/tei:title[@type eq 'main']"> a main
title must be supplied</sch:assert>
    </sch:rule>
    <sch:rule context="tei:teiHeader[ not( $uses_old_encoding ) ]">
      <sch:assert test="tei:fileDesc/tei:titleStmt/tei:title[ not( @type eq 'sub' ) ]">
a main title must be supplied</sch:assert>
    </sch:rule>
  </constraint>
</constraintSpec>
```

As a further example, Schematron may be used to enforce rules applicable to a TEI document which is going to be rendered into accessible HTML, for example to check that some sort of content is available from which the *alt* attribute of an HTML `<img>` can be created:

```
<constraintSpec ident="alt"
  scheme="schematron"
   xmlns:sch="http://purl.oclc.org/dsdl/schematron">
  <constraint>
    <sch:pattern id="altTags">
```

```
     <sch:rule  context="tei:figure">
       <sch:assert  test="tei:figDesc or tei:head"> You should provide information in a
figure from
               which we can construct an alt attribute in HTML </sch:assert>
     </sch:rule>
    </sch:pattern>
  </constraint>
</constraintSpec>
```

Schematron rules can also be used to enforce other HTML accessibility rules about tables; note here the use of a report and an assertion within one pattern:

```
<constraintSpec  ident="tables"
  scheme="schematron"
   xmlns:sch="http://purl.oclc.org/dsdl/schematron">
  <constraint>
    <sch:pattern  id="Tables">
      <sch:rule  context="tei:table">
        <sch:assert  test="tei:head">A <table> should have a caption, using a <head>
element</sch:assert>
       <sch:report  test="parent::tei:body">Do not use tables to lay out the document
body</sch:report>
      </sch:rule>
    </sch:pattern>
  </constraint>
</constraintSpec>
```

Constraints can be expressed using any convenient language. The following example uses a pattern matching language called SPITBOL to express the requirement that title and author should be different.

```
<constraintSpec  ident="title_ne_author"
  scheme="SPITBOL">
  <constraint> (output = leq(title,author) "title and author cannot be the same")
</constraint>
</constraintSpec>
```

Note that the value of *scheme* is SPITBOL. In order to properly constrain and document the values of *scheme* used in their customization file, a project may wish to create a customization that (among other things) adds and explains this value for use in validating their customization file. Thus using schemes other than those provided for by the TEI (currently only schematron) may require somewhat more effort when creating a customization file. Such private schemes will generally be even more problematic on implementation of the constraints themselves, as it may require significant programming work. The TEI only provides this capability for the suggested value (schematron).

### 23.5.3 Attribute List Specification

The <attList> element is used to document information about a collection of attributes, either within an <elementSpec>, or within a <classSpec>. An attribute list can be organized either as a group of attribute definitions, all of which are understood to be available, or as a choice of attribute definitions, of which only one is understood to be available. An attribute list may thus contain nested attribute lists.

The attribute *org* is used to indicate whether its child <attDef> elements are all to be made available, or whether only one of them may be used. For example, the attribute list for the element <moduleRef> contains a nested attribute list to indicate that either the *include* or the *except* attribute may be supplied, but not both:

```
<attList>
<!-- other attribute definitions here -->
  <attList org="choice">
    <attDef ident="include">
<!-- definition for the include attribute -->
    </attDef>
    <attDef ident="except">
<!-- definition for the except attribute -->
    </attDef>
  </attList>
</attList>
```

The <attDef> element is used to document a single attribute, using an appropriate selection from the common elements already mentioned and the following :

`<attDef>` (attribute definition) contains the definition of a single attribute.

`@usage` specifies the optionality of the attribute.

`<datatype>` (datatype) specifies the declared value for an attribute, by referring to any datatype defined by the chosen schema language.

`@minOccurs` (minimum number of occurrences) indicates the minimum number of times this datatype may occur in an instance of the attribute being defined.

`@maxOccurs` (maximum number of occurrences) indicates the maximum number of times this datatype may occur in an instance of the attribute being defined.

`<dataRef>` identifies the datatype of an attribute value, either by referencing an item in an externally defined datatype library, or by pointing to a TEI-defined data specification

`<defaultVal>` (default value) specifies the default declared value for an attribute.

`<valDesc>` (value description) specifies any semantic or syntactic constraint on the value that an attribute may take, additional to the information carried by the <datatype> element.

`<valList>` (value list) contains one or more <valItem> elements defining possible values.

`<valItem>` documents a single value in a predefined list of values.

The <attList> within an <elementSpec> is used to specify only the attributes which are specific to that particular element. Instances of the element may carry other attributes which are declared by the classes of which the element is a member. These extra attributes, which are shared by other elements, or by all elements, are specified by an <attList> contained within a <classSpec> element, as described in section *23.6. Class Specifications* below.

#### 23.5.3.1 Datatypes

The 'datatype' (i.e. the kind of value) for an attribute may be specified using the elements <datatype> and <dataRef>. A datatype may be defined in any of the following three ways:

- by reference to an existing TEI datatype definition, itself defined by a <dataSpec>;
- by use of its name in XML Schema Part 2: Datatypes Second Edition, the widely used datatype library maintained by the W3C as part of the definition of its schema language;
- by referencing its URI within some other datatype library.

The TEI defines a number of datatypes, each with an identifier beginning `teidata.`, which are used in preference to the datatypes available natively from a target schema such as RELAX NG or W3C Schema since the facilities provided by different schema languages vary so widely. The TEI datatypes available are described in section *1.4.2. Datatype Specifications* above. Note that each is, of necessity, mapped eventually to an externally defined datatype such as W3C Schema's `text` or `name`, possibly combined to give more expressivity, or constrained to a particular defined usage.

It is possible to reference a W3C schema datatype directly using *name*. In this case, the child <dataFacet> can be used instead of *restriction* to set W3C schema compliant restrictions on the datatype. A <dataFacet> is particularly useful for restrictions that can be difficult to impose and to read as a regular expression pattern.

```
<dataRef  name="decimal">
  <dataFacet  name="maxInclusive"
    value="360.0"/>
  <dataFacet  name="minInclusive"
    value="-360.0"/>
</dataRef>
```

Note that restrictions are either expressed with *restriction* or <dataFacet>, never both.

A datatype may be used to constrain the textual content of an element, rather than the value of an attribute. But because they are intended for use in defining ranges of attribute values, datatypes may not contain elements or attributes.

The attributes *minOccurs* and *maxOccurs* are available for the case where an attribute may take more than one value of the type specified. For example, the *columns* attribute of the <layout> element can have one or two non-negative integers as its value:

```
<attDef  ident="columns">
  <gloss  xml:lang="en"
    versionDate="2007-06-12">columns</gloss>
  <desc  versionDate="2005-01-14"
    xml:lang="en">specifies the number of columns per page</desc>
  <datatype  minOccurs="1"  maxOccurs="2">
    <dataRef  key="teidata.count"/>
  </datatype>
  <remarks  xml:lang="en"
    versionDate="2017-07-09">
    <p>If a single number is given, all pages referenced
        have this number of columns. If two numbers are given,
        the number of columns per page varies between the
        values supplied. Where <att>columns</att> is omitted
        the number is assumed to be <val>1</val>.</p>
  </remarks>
</attDef>
```

indicating that the *target* attribute may take any number of values, each being of the same datatype, namely the TEI data specification `teidata.pointer`. As is usual in XML, multiple values for a single attribute are separated by one or more white space characters. Hence, values such as `#a #b #c` or `http://example.org http://www.tei-c.org/index.xml` may be supplied.

#### 23.5.3.2 Value Specification

The <valDesc> element may be used to describe constraints on data content in an informal way: for example

```
<valDesc>must point to another <gi>align</gi>
element logically preceding this one.</valDesc>
```

```
<valDesc>Values should be Library of Congress
subject headings.</valDesc>
```

```
<valDesc>A bookseller's surname, taken from the list
in <title>Pollard and Redgrave</title>
</valDesc>
```

Constraints expressed in this way are purely documentary; to enforce them, the <constraintSpec> element described in section *23.5.2. Additional Constraints* must be used. For example, to specify that an imaginary attribute *ageAtDeath* must take positive integer values less than 150, the datatype `teidata.numeric` might be used in combination with a <constraintSpec> such as the following:

```
<attDef ident="ageAtDeath"
   xmlns:sch="http://purl.oclc.org/dsdl/schematron">
  <desc>age in years at death</desc>
  <datatype>
    <dataRef key="teidata.count"/>
  </datatype>
  <constraintSpec ident="lessThan150"
    scheme="schematron">
    <constraint>
      <sch:rule context="@ageAtDeath">
        <sch:assert test=". le 150">age at death must be an integer less than
150</sch:assert>
      </sch:rule>
    </constraint>
  </constraintSpec>
</attDef>
```

The elements <altIdent>, <equiv>, <gloss> and <desc> may all be used in the same way as they are elsewhere to describe fully the meaning of a coded value, as in the following example:

```
<valItem ident="dub">
  <altIdent xml:lang="fr">dou</altIdent>
  <equiv name="unknown"/>
  <gloss>dubious</gloss>
  <desc>used when the application of this element is doubtful or uncertain</desc>
</valItem>
```

Where all the possible values for an attribute can be enumerated, the datatype `teidata.enumerated` should be used, together with a <valList> element specifying the values and their significance, as in the following example:

```
<valList type="closed">
  <valItem ident="req">
    <gloss>required</gloss>
  </valItem>
  <valItem ident="rec">
    <gloss>recommended</gloss>
  </valItem>
  <valItem ident="opt">
    <gloss>optional</gloss>
  </valItem>
</valList>
```

Note the use of the <gloss> element here to explain the otherwise less than obvious meaning of the codes used for these values. Since this value list specifies that it is of type closed, only the values enumerated are legal, and an ODD processor will typically enforce these constraints in the schema fragment generated.

The <valList> element can also be used to provide illustrative examples of the kinds of values expected without listing all of them. In such cases the *type* attribute will have the value open, as in the following example:

```
<attDef ident="type" usage="opt">
  <desc versionDate="2005-01-14"
    xml:lang="en">characterizes the movement, for example as an
     entrance or exit.</desc>
  <desc versionDate="2007-12-20"
    xml:lang="ko">예를 들어 입장 또는 퇴장과 같은, 이동의 특성을 기술한다.</desc>
  <datatype>
    <dataRef key="teidata.enumerated"/>
  </datatype>
  <valList type="open">
    <valItem ident="entrance">
      <desc versionDate="2007-06-27"
        xml:lang="en">character is entering the stage.</desc>
      <desc versionDate="2007-12-20"
        xml:lang="ko">등장인물이 무대에 등장하고 있다.</desc>
    </valItem>
    <valItem ident="exit">
      <desc versionDate="2007-06-27"
        xml:lang="en">character is exiting the stage.</desc>
      <desc versionDate="2007-12-20"
        xml:lang="ko">등장인물이 무대에서 퇴장하고 있다.</desc>
    </valItem>
    <valItem ident="onStage">
      <desc versionDate="2007-07-04"
        xml:lang="en">character moves on stage</desc>
      <desc versionDate="2007-12-20"
        xml:lang="ko">등장인물이 무대에서 이동한다.</desc>
    </valItem>
  </valList>
</attDef>
```

The datatype will be `teidata.enumerated` in either case.

The <valList> or <dataRef> elements may also be used (as a child of the <content> element) to put constraints on the permitted content of an element, as noted at *23.5.1.2. Defining Content Models: RELAX NG*. This use is not however supported by all schema languages, and is therefore not recommended if support for non-RELAX NG systems is a consideration.

### 23.5.4 Processing Models

As far as possible, the TEI defines elements and their attributes in a way which is entirely independent of their subsequent processing, since its intention is to maximize the reusability of encoded documents and their use in multiple contexts. Nevertheless, it can be very useful to specify one or more possible models for such processing, both to clarify the intentions of the encoder, and to provide default behaviours for a software engineer to implement when documents conforming to a particular TEI customization are processed. To that end, the following elements may be used to document one or more *processing models* for a given element.

`<model>` describes the processing intended for a specified element.

`<modelGrp>` (model group) any grouping of <model> or <modelSequence> elements with a common output method.

`<modelSequence>` any sequence of model or <modelSequence> elements which is to be processed as a single set of actions.

One or more of these elements may appear directly within an element specification to define the processing anticipated for that element, more specifically how it should be processed to produce the kind of output indicated by the *output* attribute. Where multiple such elements appear directly within an <elementSpec>, they are understood to document mutually exclusive processing models, possibly for different outputs or applicable in different contexts. Alternatively, the <modelGrp> element may be used to group alternative <model> elements intended for a single

kind of output. The <modelSequence> element is provided for the case where a sequence of models is to be processed, functioning as a single unit.

A processing model suggests how a given element may be transformed to produce one or more outputs. The model is expressed in terms of *behaviours* and *parameters*, using high-level formatting concepts familiar to software engineers and web designers, such as 'block' or 'inline'. As such, it has a different purpose from existing TEI mechanisms for documenting the appearance of source materials, such as the global attributes *rend*, *rendition* and *style*, described in sections *2.3.4.1. Rendition* and *3.3.1. What Is Highlighting?*. It does not necessarily describe anything present in the original source, nor does it necessarily represent its original structure or semantics. A processing model is a template description, which may be used to simplify the task of producing or customizing the stylesheets needed by a formatting engine or any other form of processor.

#### 23.5.4.1 The TEI processing model

The <model> element is used to document the processing model intended for a particular element in an abstract manner, independently of its implementation in whatever processing language is chosen. This is achieved by means of the following attributes and elements:

`<model>` describes the processing intended for a specified element.

- `@predicate [att.predicate]` the condition under which the element bearing this attribute applies, given as an XPath predicate expression.
- `@behaviour` names the process or function which this processing model uses in order to produce output.
- `@output` the intended output.
- `@useSourceRendition` whether to obey any rendition attribute that is present.
- `@cssClass` the name of a CSS class which should be associated with this element

`<outputRendition>` describes the rendering or appearance intended for all occurrences of an element in a specified context for a specified type of output.

- `@scope` provides a way of defining 'pseudo-elements', that is, styling rules applicable to specific sub-portions of an element.

The mandatory *behaviour* attribute defines in broad terms how an element should be processed, for example as a block or as an inline element. The optional *predicate* attribute may be used to specify a subset of contexts in which this model should be applicable: for example, an element might be treated as a block element in some contexts, but not in others. The *output* attribute supplies a name for the output for which this model is intended, for example for screen display, for a printed reading copy, for a scholarly publication, etc. The way in which an element should be rendered is declared independently of its behaviour, using either the attribute *useSourceRendition* or the element <outputRendition>. These Guidelines recommend that <outputRendition> be expressed using the W3C Cascading Stylesheet Language (CSS), but other possibilities are not excluded. The particular language used may be documented by means of the <styleDefDecl> element described in *2.3.5. The Default Style Definition Language Declaration*.

#### 23.5.4.2 Output Rendition

Here is a simple example of a processing model which might be included in the specification for an element such as <hi> or <foreign>. The intent is that these elements should be presented inline using an italic font.

```
<model behaviour="inline">
  <outputRendition>font-style: italic;</outputRendition>
</model>
```

If the <rendition> element, or the attributes *style*, *rend*, or *rendition* have already been used in the source document to indicate elements that were originally rendered in italic, and we wish simply to follow this in our processing, then there is no need to include an <outputRendition> element, and the attribute *useSourceRendition* could be used as follows:

```
<model behaviour="inline"
  useSourceRendition="true"/>
```

Any rendition information present in the source document will be ignored unless the *useSourceRendition* attribute has the value true. If that is the case, then such information will be combined with any rendition information supplied by the <outputRendition> element. For example, using CSS, an element which appears in the source as follows

```
<hi  style="font-weight:bold;">this is in bold</hi>
```

would appear in bold and italic if processed by the following model

```
<model  behaviour="inline"
  useSourceRendition="true">
  <outputRendition>font-style: italic;</outputRendition>
</model>
```

#### 23.5.4.3 CSS Class

In a typical workflow processing TEI documents for display on the web, a system designer will often wish to use an externally defined CSS stylesheet. The *cssClass* attribute simplifies the task of maintaining compatibility amongst the possibly many applications using such a stylesheet and also enables a TEI application to specify the names of classes to be used for particular processing models. For example, supposing that the associated CSS stylesheet includes a CSS class called `labeled-list`, the following processing model might be used to request it be used for <list> elements containing a child <label> element:

```
<elementSpec  ident="list"  mode="change">
  <model  predicate="label"  behaviour="list"
    cssClass="labeled-list">
<!-- ... -->
  </model>
</elementSpec>
```

In the following example, a table will be formatted using renditional information provided in the source if that is available, or by an external stylesheet, using one of the CSS classes specified, if it is not:

```
<elementSpec  mode="change"  ident="table">
<!-- Preserve original rendition for tables which contain @rendition hints -->
  <model  predicate=".//row/@rendition or .//cell/@rendition"
    behaviour="table"  useSourceRendition="true"/>
<!-- Use bootstrap for default table styling -->
  <model  behaviour="table"
    useSourceRendition="true"
    cssClass="table table-hover table-bordered"/>
</elementSpec>
```

As discussed further below, the input data available to a processing model is by default the content of the element being processed, together with its child nodes.

#### 23.5.4.4 Model Contexts and Outputs

Sometimes different processing models are required for the same element in different contexts. For example, we may wish to process the <quote> element as an inline italic element when it appears inside a <p> element, but as an indented block when it appears elsewhere. To achieve this, we need to change the specification for the <quote> element to include two <model> elements as follows:

```
<elementSpec ident="quote" mode="change">
  <model predicate="ancestor::p"
    behaviour="inline">
    <outputRendition>font-style: italic;</outputRendition>
  </model>
  <model behaviour="block">
    <outputRendition>left-margin: 2em;</outputRendition>
  </model>
</elementSpec>
```

As noted above, these two models are mutually exclusive. The first processing model will be used only for <quote> elements which match the XPath expression given as value for the *predicate* attribute. All other element occurrences will use the second processing model.

When, as here, multiple behaviours are required for the same element, it will often be the case that the appropriate processing will depend on the context. It may however be the case that the choice of an appropriate model will be made on the basis of the intended output. For example, we might wish to define quite different behaviours when a document is to be displayed on a mobile device and when it is to be displayed on a desktop screen. Different behaviours again might be specified for a print version intended for the general reader, and for a print version aimed at the technical specialist.

The <modelGrp> element can be used to group together all the processing models which have in common a particular intended output, as in the following example:

```
<modelGrp output="mobile">
  <model behaviour="inline"
    predicate="@rend='inline'">
    <outputRendition>font-size: 7pt;</outputRendition>
  </model>
  <model behaviour="block"
    predicate="@rend='block'">
    <outputRendition>text-color: red;</outputRendition>
  </model>
</modelGrp>
<modelGrp output="print">
  <model behaviour="inline"
    predicate="@rend='inline'">
    <outputRendition>font-size: 12pt;</outputRendition>
  </model>
  <model behaviour="block"
    useSourceRendition="true" predicate="@rend='block'">
    <outputRendition>text-align: center;</outputRendition>
  </model>
</modelGrp>
```

#### 23.5.4.5 Behaviours and their parameters

In the examples above we have used without explanation or definition two simple *behaviours*: `inline` and `block`, but many other behaviours are possible. A list of recommended behaviour names forms part of the specification for the element <model>. A processing model can specify any named behaviour, some of which have additional *parameters*. The `parameters` of a `behaviour` resemble the arguments of a function in many programming languages: they provide names which can be used to distinguish particular parts of the input data available to the process used to implement the behaviour in question.

The following elements are used to represent and to define parameters:

`<param>` provides a parameter for a model behaviour by supplying its name and an XPath expression identifying the location of its content.

`@name` a name for the parameter being supplied.

`<paramList>` list of parameter specifications.

`<paramSpec>` supplies specification for one parameter of a model behaviour.

By default, a processor implementing the TEI processing model for a particular element has available to it as input data the content of the element itself, and all of its children. One or more <param> elements may be supplied within a <model> element to specify *parameters* which modify this, either by selecting particular parts of the default input data, or by selecting data which would not otherwise be available. In either case, the value supplied for the parameter is given as an XPath expression, evaluated with respect to the element node being processed. An arbitrary name (defined in the corresponding <paramSpec>) is also supplied to a processor to identify each parameter.

For example, an element such as the TEI <ref> element will probably be associated with a processing model which treats it as a hyperlink. But a hyperlink (in most implementations) often has two associated pieces of information: the address indicated, and some textual content serving to label the link . In HTML, the former is provided as value of the *href* element, and the latter by the content of an `<a>` element. In the following processing model we define a behaviour called `link`, which will use whatever is indicated by the parameter called `uri` to provide the former, while the latter is provided by the content of the <ref> element itself:

```
<elementSpec ident="ref" mode="add">
  <model behaviour="link">
    <param name="uri" value="@target"/>
    <param name="content" value="."/>
  </model>
</elementSpec>
```

The *value* attribute of a <param> element supplies an XPath expression that indicates where the required value may be found. The context for this XPath is the element which is being processed; hence in this example, the `uri` parameter takes the value of the *target* attribute on the <ref> element being processed. The `content` parameter indicates that the content of that <ref> element should be provided as its value. (This parameter is not strictly necessary, since by default the whole content of the element being processed is always available to a processor, but supplying it in this way makes the procedure more explicit).

All the parameters available for a given behaviour are defined as a part of the definition of the behaviour itself, as further discussed in section *23.5.4.8. Defining a processing model* below.

As a further example, the TEI <choice> element requires a different behaviour for which the name `alternate` is proposed as in the following example:

```
<elementSpec ident="choice" mode="change">
  <model predicate="sic and corr"
    behaviour="alternate">
    <param name="default" value="corr"/>
    <param name="alternate" value="sic"/>
  </model>
</elementSpec>
```

The processing model shown here will be selected for processing a <choice> element which has both <sic> and <corr> child elements. The names `default` and `alternate` here are provided for convenience. The `default` parameter provides the value of the child <corr> element, and the `alternate` parameter will provide that of the child <sic> elements. If neither <param> element was supplied, both elements would still be available to an application, but the application would need to distinguish them for itself.

A <choice> element might contain multiple corrections, each with differing values for their *cert* attribute. In the following processing model, we will accept as value of the `default` attribute only those child <corr> elements which have a value high for that attribute:

```
<elementSpec ident="choice" mode="change">
  <model predicate="sic and corr"
```

```
    behaviour="alternate">
    <param  name="default"
      value="corr[@cert='high']"/>
    <param  name="alternate"  value="sic"/>
  </model>
</elementSpec>
```

A <choice> element might contain several different pairs of alternate elements (<abbr> and <expan>, <orig> and <reg>, etc.) We might wish to group together a set of processing models for these, for example to determine which of the possible alternatives is displayed by default whenever a <choice> element is processed for output to the web:

```
<elementSpec  ident="choice"  mode="change">
  <modelGrp  output="web">
    <model  predicate="sic and corr"
      behaviour="alternate">
      <param  name="default"
        value="corr[@cert='high']"/>
      <param  name="alternate"  value="sic"/>
    </model>
    <model  predicate="abbr and expan"
      behaviour="alternate">
      <param  name="default"  value="expan[1]"/>
      <param  name="alternate"  value="abbr"/>
    </model>
    <model  predicate="orig and reg"
      behaviour="alternate">
      <param  name="default"  value="reg"/>
      <param  name="alternate"  value="orig"/>
    </model>
  </modelGrp>
</elementSpec>
```

If nothing matches the XPath defining the value of a particular parameter (e.g. if in the above example there is no correction with *cert*=high) then the `default` parameter has no value. It is left to implementors to determine how null-valued parameters should be processed.

#### 23.5.4.6 Outputs

As noted above, the *output* attribute is used to associate particular processing models with a specific type of output. The following example documents a range of processing intentions for the <date> element, intended to cope with at least the following three situations:

1. there is text inside the element, and the output is print;
2. there is no text inside the element but there is a *when* attribute, and the output is print;
3. there is a *when* attribute, there is text inside the element, and the output is web

```
<elementSpec  ident="date"  mode="change">
  <modelGrp  output="print">
    <model  predicate="text()"
      behaviour="inline"/>
    <model  predicate="@when and not(text())"
      behaviour="inline">
      <param  name="content"  value="@when"/>
    </model>
```

```
  </modelGrp>
  <model output="web" predicate="@when"
    behaviour="alternate">
    <param name="default" value="."/>
    <param name="alternate" value="@when"/>
  </model>
</elementSpec>
```

For output to print we supply two processing models, one for the simplest case where the content of the <date> is to be treated as an inline element, and the other for the case where there is no content and the value of the *when* attribute is to be used in its place. This is specified by a parameter, called `content` in this example. For output to web, we use the `alternate` behaviour discussed in the previous section to indicate that by default the content of the element will be used, while retaining access to the value of the *when* attribute, this time via a parameter called `alternate`.

#### 23.5.4.7 Model sequence

As well as being combined to form model groups, several models may be combined to form a *model sequence*. All of the individual components of a model sequence are understood to be applied, rather than considered to be mutually exclusive alternatives. For example, we might wish to define two different behaviours for a <note> element: the inline behaviour should be used to display the value of the note number (given by its *n* attribute), while a different behaviour (here called `footnote`) should be used to display the content of the element at a specified place, given by the parameter `place`. Because both of these actions are required, the two models are grouped by a <modelSequence> element:

```
<elementSpec ident="note" mode="change">
  <modelSequence output="print">
    <model behaviour="inline">
      <param name="content" value="@n"/>
    </model>
    <model behaviour="footnote">
      <param name="place" value="'foot'"/>
    </model>
  </modelSequence>
</elementSpec>
```

The value of the parameter called `place` above is an XPath expression supplying an arbitrary string ('foot'), which is therefore quoted. It is left to implementors to validate or constrain the possible values for such expressions.

#### 23.5.4.8 Defining a processing model

The processing model for an element is defined using some combination of <model>, <modelSequence>, or <modelGrp> elements within the <elementSpec> element containing its specification. The processing to be carried out is defined by means of the behaviour specified for each <model> element, possibly supplying specified values for a number of named parameters. The parameters available for a given behaviour are specified using a number of <param> elements grouped together in a <paramList> element. This <paramList> is supplied within the <valItem> used to document and name the behaviour. Here for example is the <valItem> which defines the `link` behaviour presented above:

```
<valItem ident="link">
  <desc>create a hyperlink</desc>
  <paramList>
    <paramSpec ident="content">
      <desc>supplies the location of some content describing the link</desc>
    </paramSpec>
```

```
    <paramSpec ident="uri">
      <desc>supplies the location of the intended hyperlink</desc>
    </paramSpec>
  </paramList>
</valItem>
```

Similarly the <valItem> which defines the behaviour named `alternate` includes specifications for two parameters: one also called `alternate` and the other called `default`

```
<valItem ident="alternate">
  <desc versionDate="2015-08-21"
    xml:lang="en">support display of alternative visualisations, for
     example by displaying the preferred content, by displaying both in parallel, or by
toggling
     between the two.</desc>
  <paramList>
    <paramSpec ident="default">
      <desc versionDate="2015-08-21"
        xml:lang="en">supplies the location of the preferred
            content</desc>
    </paramSpec>
    <paramSpec ident="alternate">
      <desc versionDate="2015-08-21"
        xml:lang="en">supplies the location of the alternative
            content</desc>
    </paramSpec>
  </paramList>
</valItem>
```

The suggested behaviours provided by the <model> element are informally defined using commonly understood terminology, but specific details of how they should be implemented are left to the implementor. Such decisions may vary greatly depending on the kind of processing environment, the kind of output envisaged, etc. The intention is to reduce as far as possible any requirement for the implementor to be aware of TEI-specific rules, and to maximize the ability of the ODD to express processing intentions without fully specifying an implementation.

#### 23.5.4.9 Implementation of Processing Models

As elsewhere in these Guidelines, the recommendations of this section are intended to be implementation-agnostic, not favouring any particular implementation method or technology. An implementor may choose, for example, whether to pre-process processing model specifications into a free standing 'pipeline', or to extract and process them dynamically during document processing. However, although implementors are generally free to interpret the processing model documentation according to their own requirements, some general assumptions underlie the recommendations made here:

- If a <model> element has no child <param> elements, the action specified by its behaviour should be applied to the whole element node concerned, including its child nodes of whatever type. If that behaviour requires distinguishing particular parts of the input, an implementation may choose either to make those distinctions itself, or to raise an error.
- If a <model> element has no *predicate* or *output* attribute then it is assumed to apply to all instances of the element defined in its parent <elementSpec> for all outputs. Otherwise its applicability depends on these attributes.
- Only one of the <model>s is to be applied for a particular instance of the element, except when they appear within a <modelSequence>
- A 'matching' model is one where the element to be processed satisfies the XPath in the *predicate* attribute of the <model> or <modelSequence> and the current output method matches the method specified in the *output* attribute of the <model>, <modelSequence>, or a containing <modelGrp>. A <model> or <modelSequence> without a

*predicate* always matches the element to be processed. A <model>, <modelGrp>, or <modelSequence> without an *output* attribute matches any output method.

- Processing Model implementations must execute only the first matching <model> or <modelSequence> in document order.
- If there are two or more <model> elements provided for an <elementSpec> but they have different *output* attributes then the implementation should choose the <model> appropriate to the desired output.
- If there are two or more <model> elements provided for an <elementSpec> but they have different *predicate* attributes then the implementation should choose the <model> whose *predicate* provides the most specific context (where *specific* is understood in the same way as in XSLT)

In the following example, which shows part of the element specification for the <head> element, there are two <model> elements, one with and one without a *predicate* attribute:

```
<model behaviour="inline"
  predicate="parent::list">
  <desc versionDate="2015-03-02"
    xml:lang="en">Model for list headings</desc>
<!-- ... -->
</model>
<model behaviour="heading">
  <desc versionDate="2016-03-02"
    xml:lang="en">Default model for all headings.</desc>
<!-- ... -->
</model>
```

In this example, an implementation should use the first processing model only for <head> elements with a <list> element as parent; for all other <head> elements, the second processing model should be used.

## 23.6 Class Specifications

The element <classSpec> is used to document either an *attribute class* or a 'model class', as defined in section *1.3. The TEI Class System*. A corresponding <classRef> element may be used to select a specific named class from those available.

`<classSpec>` (class specification) contains reference information for a TEI element class; that is a group of elements which appear together in content models, or which share some common attribute, or both.
`@type` indicates whether this is a model class or an attribute class

`<classRef>` points to the specification for an attribute or model class which is to be included in a schema.
`@expand` indicates how references to this class within a content model should be interpreted.

`<attList>` (attribute list) contains documentation for all the attributes associated with this element, as a series of <attDef> elements.
`@org` (organization) specifies whether only one (choice) or all (group) of the attributes in the list are available.

A model class specification does not list all of its members. Instead, its members declare that they belong to it by means of a <classes> element contained within the relevant <elementSpec>. This will contain a <memberOf> element for each class of which the relevant element is a member, supplying the name of the relevant class. For example, the <elementSpec> for the element <hi> contains the following:

```
<classes>
  <memberOf key="model.hiLike"/>
</classes>
```

This indicates that the <hi> element is a member of the class with identifier `model.hiLike`. The <classSpec> element that documents this class contains the following declarations:

```
<classSpec  type="model"
  ident="model.hiLike">
  <desc>groups phrase-level elements related to highlighting that have no specific
semantics </desc>
  <classes>
    <memberOf  key="model.highlighted"/>
  </classes>
</classSpec>
```

which indicate that the class `model.hiLike` is actually a member (or subclass) of the class `model.highlighted`.

The function of a model class declaration is to provide another way of referring to a group of elements. It does not confer any other properties on the elements which constitute its membership.

The attribute *type* is used to distinguish between 'model' and 'attribute' classes. In the case of attribute classes, the attributes provided by membership in the class are documented by an <attList> element contained within the <classSpec>. In the case of model classes, no further information is needed to define the class beyond its description, its identifier, and optionally any classes of which it is a member.

When a model class is referenced in the content model of an element (i.e. by means of a <classRef> element within the <content> of an <elementSpec>), its meaning will depend on the value of its *expand* attribute.

If this attribute is not specified, the <classRef> is interpreted to mean an alternated list of all the current members of the class named. For example, suppose that the members of the class `model.hiLike` are elements <hi>, `<it>`, and `<bo>`. Then a content model such as

```
<content>
  <classRef  key="model.hiLike"/>
</content>
```

would be equivalent to the explicit content model:

```
<content>
  <alternate>
    <elementRef  key="hi"/>
    <elementRef  key="it"/>
    <elementRef  key="bo"/>
  </alternate>
</content>
```

(or, to use RELAX NG compact syntax, `( hi | it | bo )`). However, a content model of `<classRef expand="sequence"/>` would be equivalent to the following explicit content model:

```
<content>
  <sequence>
    <elementRef  key="hi"/>
    <elementRef  key="it"/>
    <elementRef  key="bo"/>
  </sequence>
</content>
```

(or, in RELAX NG compact syntax, `( hi, it, bo )`).

An attribute class (a <classSpec> of *type* atts) contains an <attList> element which lists the attributes that all the members of that class inherit from it. For example, the class `att.interpLike` defines a small set of attributes common to all elements which are members of that class: those attributes are listed by the <attList> element

contained by the <classSpec> for `att.interpLike`. When processing the documentation elements for elements which are members of that class, an ODD processor is required to extend the <attList> (or equivalent) for such elements to include any attributes defined by the <classSpec> elements concerned. There is a single global attribute class, `att.global`, to which some modules contribute additional attributes when they are included in a schema.

## 23.7 Macro Specifications

The <macroSpec> element is used to declare and document predefined strings or patterns not otherwise documented by the elements described in this section. A corresponding <macroRef> element may be used to select a specific named pattern from those available. Patterns are used as a shorthand chiefly to describe common content models and datatypes, but may be used for any purpose. The following elements are used to represent patterns:

`<macroSpec>` (macro specification) documents the function and implementation of a pattern.

`<macroRef>` points to the specification for some pattern which is to be included in a schema.

- `@key` the identifier used for the required pattern within the source indicated.

## 23.8 Building a TEI Schema

The specification elements, and some of their children, are all members of the `att.identified` class, from which they inherit the following attributes:

`att.identified` provides identifying attributes for elements which can be subsequently referenced by means of a *key* attribute.

- `@ident` supplies the identifier by which this element may be referenced.
- `@predeclare` says whether this object should be predeclared in the `tei` infrastructure module.
- `@module` supplies a name for the module in which this object is to be declared.

This attribute class is a subclass of the `att.combinable` class from which it (and some other elements) inherits the following attribute:

`att.combinable` provides attributes indicating how multiple references to the same object in a schema should be combined.

- `@mode` specifies the effect of this declaration on its parent object.

This attribute class, in turn, is a subclass of the `att.deprecated` class, from which it inherits the following attribute:

`att.deprecated` provides attributes indicating how a deprecated feature will be treated in future releases.

- `@validUntil` provides a date before which the construct being defined will not be removed.

The *validUntil* attribute may be used to signal an intent to remove a construct from future versions of the schema being specified.

The <elementSpec>, <attDef> and <schemaSpec> specification elements also have an attribute which determines which namespace to which the object being created will belong. In the case of <schemaSpec>, this namespace is inherited by all the elements created in the schema, unless they have their own *ns*.

`att.namespaceable` provides attributes indicating the target namespace for an object being created.

These attributes are used by an ODD processor to determine how declarations are to be combined to form a schema or DTD, as further discussed in this section.

### 23.8.1 TEI customizations

As noted above, a TEI schema is defined by a <schemaSpec> element containing an arbitrary mixture of explicit declarations for objects (i.e. elements, classes, patterns, or macro specifications) and references to other objects containing such declarations (i.e. references to specification groups, or to modules). A major purpose of this mechanism is to simplify the process of defining user customizations, by providing a formal method for the user to combine new declarations with existing ones, or to modify particular parts of existing declarations.

In the simplest case, a user-defined schema might just combine all the declarations from two nominated modules:

```
<schemaSpec ident="example">
  <moduleRef key="core"/>
```

```
  <moduleRef key="linking"/>
</schemaSpec>
```

An ODD processor, given such a document, should combine the declarations which belong to the named modules, and deliver the result as a schema of the requested type. It may also generate documentation for the elements declared by those modules. No source is specified for the modules, and the schema will therefore combine the declarations found in the most recent release version of the TEI Guidelines known to the ODD processor in use.

The value specified for the *source* attribute, when it is supplied as a URL, specifies any convenient location from which the relevant ODD files may be obtained. For the current release of the TEI Guidelines, a URL in the form `http://www.tei-c.org/Vault/P5/x.y.z/xml/tei/odd/p5subset.xml` may be used, where `x.y.z` represents the P5 version number, e.g. `1.3.0`. Alternatively, if the ODD files are locally installed, it may be more convenient to supply a value such as ../ODDs/p5subset.xml".

The value for the *source* attribute may be any form of URI. A set of TEI-conformant specifications in a form directly usable by an ODD processor must be available at the location indicated. When no *source* value is supplied, an ODD processor may either raise an error or assume that the location of the current release of the TEI Guidelines is intended.

If the source is specified in the form of a private URI, the form recommended is `aaa:x.y.z`, where `aaa` is a prefix indicating the markup language in use, and `x.y.z` indicates the version number. For example, `tei:1.2.1` should be used to reference release 1.2.1 of the current TEI Guidelines. When such a URI is used, it will usually be necessary to translate it before such a file can be used in blind interchange.

The effect of a <moduleRef> element is to include in the schema all declarations provided by that module. This may be modified by means of the attributes *include* and *except* which allow the encoder to supply an explicit lists of elements from the stated module which are to be included or excluded respectively. For example:

```
<schemaSpec ident="example">
  <moduleRef key="core"
    except="add del orig reg"/>
  <moduleRef key="linking"
    include="linkGrp link"/>
</schemaSpec>
```

The schema specified here will include all the elements supplied by the core module except for <add>, <del>, <orig>, and <reg>. It will also include only the <linkGrp> and <link> elements from the linking module.

Alternatively, the element <elementRef> may be used to indicate explicitly which elements are to be included in a schema. The same effect as the preceding example might thus be achieved by the following:

```
<schemaSpec ident="example">
  <moduleRef key="core"
    except="add del orig reg"/>
  <elementRef key="linkGrp"/>
  <elementRef key="link"/>
</schemaSpec>
```

Note that in this last case, there is no need to specify the name of the module from which the two element declarations are to be found; in the TEI scheme, element names are unique across all modules. The module is simply a convenient way of grouping together a number of related declarations.

In the same way, a schema may select a subset of the attributes available in a specific class, using the <classRef> element to point to an attribute class:

```
<schemaSpec ident="example">
  <moduleRef key="tei"/>
<!-- ... -->
  <classRef key="att.global.linking"
    include="corresp"/>
<!-- ... -->
</schemaSpec>
```

Here, only the *corresp* attribute is included; the other attributes in the class are not available. The same effect can be achieved using *except*:

```
<schemaSpec ident="example">
  <moduleRef key="tei"/>
<!-- ... -->
  <classRef key="att.global.linking"
    except="copyOf exclude next prev sameAs select synch"/>
<!-- ... -->
</schemaSpec>
```

A schema may also include declarations for new elements, as in the following example:

```
<schemaSpec ident="example">
  <moduleRef key="header"/>
  <moduleRef key="verse"/>
  <elementSpec ident="soundClip">
    <classes>
      <memberOf key="model.pPart.data"/>
    </classes>
  </elementSpec>
</schemaSpec>
```

A declaration for the element `<soundClip>`, which is not defined in the TEI scheme, will be added to the output schema. This element will also be added to the existing TEI class `model.pPart.data`, and will thus be available in TEI conformant documents. Attributes from existing TEI classes could be added to the new element using <attRef>:

```
<schemaSpec ident="example">
  <moduleRef key="header"/>
  <moduleRef key="verse"/>
  <elementSpec ident="soundClip">
    <classes>
      <memberOf key="model.pPart.data"/>
    </classes>
    <attList>
      <attRef class="att.global.source"
        name="source"/>
    </attList>
  </elementSpec>
</schemaSpec>
```

This will provide the *source* attribute from the `att.global.source` class on the new `<soundClip>` element.

A schema might also include re-declarations of existing elements, as in the following example:

```
<schemaSpec ident="example">
  <moduleRef key="header"/>
  <moduleRef key="teistructure"/>
  <elementSpec ident="head" mode="change">
    <content>
      <macroRef key="macro.xtext"/>
    </content>
  </elementSpec>
</schemaSpec>
```

The effect of this is to redefine the content model for the element <head> as plain text, by over-riding the <content> child of the selected <elementSpec>. The attribute specification `mode="change"` has the effect of over-riding only those children elements of the <elementSpec> which appear both in the original specification and in the new specification supplied above: <content> in this example. Note that if the value for *mode* were replace, the effect would be to replace all children elements of the original specification with the children elements of the new specification, and thus (in this example) to delete all of them except <content>.

A schema may not contain more than two declarations for any given component. The value of the *mode* attribute is used to determine exactly how the second declaration (and its constituents) should be combined with the first. The following table summarizes how a processor should resolve duplicate declarations; the term *identifiable* refers to those elements which can have a *mode* attribute:

| mode value | existing declaration | effect |
|---|---|---|
| add | no | add new declaration to schema; process its children in add mode |
| add | yes | raise error |
| replace | no | raise error |
| replace | yes | retain existing declaration; process new children in replace mode; ignore existing children |
| change | no | raise error |
| change | yes | process identifiable children according to their modes; process unidentifiable children in replace mode; retain existing children where no replacement or change is provided |
| delete | no | raise error |
| delete | yes | ignore existing declaration and its children |

### 23.8.2 Combining TEI and Non-TEI Modules

In the simplest case, all that is needed to include a non-TEI module in a schema is to reference its RELAX NG source using the *url* attribute on <moduleRef>. The following specification, for example, creates a schema in which declarations from the non-TEI module `svg11.rng` (defining Standard Vector Graphics) are included. To avoid any risk of name clashes, the schema specifies that all TEI patterns generated should be prefixed by the string ”TEI_”.

```
<schemaSpec prefix="TEI_" ident="testsvg"
  start="TEI svg">
  <moduleRef key="header"/>
  <moduleRef key="core"/>
  <moduleRef key="tei"/>
  <moduleRef key="textstructure"/>
  <moduleRef url="svg11.rng"/>
</schemaSpec>
```

This specification generates a single schema which might be used to validate either a TEI document (with the root element <TEI>), or an SVG document (with a root element `<svg:svg>`), but would *not* validate a TEI document

containing <svg:svg> or other elements from the SVG language. For that to be possible, the <svg:svg> element must become a member of a TEI model class (*1.3. The TEI Class System*), so that it may be referenced by other TEI elements. To achieve this, we modify the last <moduleRef> in the above example as follows:

```
<moduleRef  url="svg11.rng">
  <content>
    <rng:define  name="TEI_model.graphicLike"
      combine="choice">
      <rng:ref  name="svg"/>
    </rng:define>
  </content>
</moduleRef>
```

This states that when the declarations from the svg11.rng module are combined with those from the other modules, the declaration for the model class model.graphicLike in the TEI module should be extended to include the element <svg:svg> as an alternative. This has the effect that elements in the TEI scheme which define their content model in terms of that element class (notably <figure>) can now include it. A RELAX NG schema generated from such a specification can be used to validate documents in which the TEI <figure> element contains any valid SVG representation of a graphic, embedded within an <svg:svg> element.

### 23.8.3 Linking Schemas to XML Documents

Schemas can be linked to XML documents by means of the <?xml-model?> processing instruction described in the W3C Working Group Note *Associating Schemas with XML documents* (https://www.w3.org/TR/xml-model/). <?xml-model?> can be used for any type of schema, and may be used for multiple schemas:

```
  <?xml-model href="tei_ tite.rng" type="application/xml" ?>
<?xml-model href="checkLinks.sch" type="application/xml" schematypens="http://purl.oclc.org/dsdl/
?>
<?xml-model href="tei_ tite.odd" type="application/tei+xml" schematypens="http://www.tei-c.org/ns
?>
```

This example includes a standard RELAX NG schema, a Schematron schema which might be used for checking that all pointing attributes point at existing targets, and also a link to the TEI ODD file from which the RELAX NG schema was generated. See also *2.3.10. The Schema Specification* for details of another method of linking an ODD specification into your file by including a <schemaSpec> element in <encodingDesc>.

## 23.9 Module for Documentation Elements

The module described in this chapter makes available the following components:

**Module tagdocs: Documentation of TEI and other XML markup languages**

- *Elements defined*: altIdent alternate anyElement att attDef attList attRef classRef classSpec classes code constraint constraintDecl constraintSpec content dataFacet dataRef dataSpec datatype defaultVal eg egXML elementRef elementSpec empty equiv exemplum gi ident listRef macroRef macroSpec memberOf model modelGrp modelSequence moduleRef moduleSpec outputRendition param paramList paramSpec remarks schemaSpec sequence specDesc specGrp specGrpRef specList tag textNode val valDesc valItem valList
- *Classes defined*: att.combinable att.deprecated att.identified att.namespaceable att.predicate att.repeatable att.translatable model.contentPart

The selection and combination of modules to form a TEI schema is described in *1.2. Defining a TEI Schema.*

The elements described in this chapter are all members of one of three classes: model.oddDecl, model.oddRef, or model.phrase.xml, with the exceptions of <schemaSpec> (a member of

`model.divPart`) and both <eg> and <egXML> (members of `model.common` and `model.egLike`). All of these classes are declared along with the other general TEI classes, in the basic structure module documented in *1. The TEI Infrastructure*.

In addition, some elements are members of the `att.identified` class, which is documented in *23.8.1. TEI customizations* above.

# 24 Using the TEI

This section discusses some technical topics concerning the deployment of the TEI markup schema documented elsewhere in these Guidelines. In section *24.3. Customization* we discuss the scope and variety of the TEI customization mechanisms. In *24.4. Conformance* we define the notions of *TEI Conformance* and *TEI Extension.* Since the ODD markup description language defined in chapter *23. Documentation Elements* is fundamental to the way conformance and customization are handled in the TEI system, these two definitional sections are followed by a section (*24.5. Implementation of an ODD System*) which describes the intended behaviour of an ODD processor.

## 24.1 Serving TEI files with the TEI Media Type

In February 2011, the media type `application/tei+xml` was registered with IANA for 'markup languages defined in accordance with the Text Encoding and Interchange guidelines' (RFC 6129). We recommend that any XML file whose root element is in the TEI namespace be served with the media type `application/tei+xml` to enable and encourage automated recognition and processing of TEI files by external applications.

## 24.2 Obtaining the TEI

As discussed in chapter *23. Documentation Elements*, all components of the TEI schema are generated from a single set of TEI XML source files. Schemas can be generated in each of XML DTD language, W3C schema language, and RELAX NG schema language. Documentation can be generated in a variety of commonly-used document formats, including HTML, DOCX, or PDF.

TEI components are freely available over the Internet and elsewhere. The canonical home for the TEI source, the schema fragments generated from it, and example modifications, is the TEI repository at `https://github.com/TEIC/TEI`; versions are also available in other formats, along with copies of these Guidelines and related materials, from the TEI web site at `http://www.tei-c.org/`.

## 24.3 Customization

These Guidelines provide an encoding schema suitable for encoding a very wide range of texts, and capable of supporting a wide variety of applications. For this reason, the TEI schema supports a variety of different approaches to solving similar problems, and also defines a much richer set of elements than is likely to be necessary in any given project. Furthermore, the TEI schema may be extended in well-defined and documented ways for texts that cannot be conveniently or appropriately encoded using what is provided. For these reasons, it is almost impossible to use the TEI schema without customizing it in some way.

This section describes how the TEI encoding schema may be customized, and should be read in conjunction with chapter *23. Documentation Elements*, which describes how a specific application of the TEI encoding schema should be documented. The documentation system described in that chapter is, like the rest of the TEI schema, independent of any particular schema or document type definition language.

Formally speaking, these Guidelines provide both syntactic rules about how elements and attributes may be used in valid documents and semantic recommendations about what interpretation should be attached to a given syntactic construct. In this sense, they provide both a *document type definition* and a *document type declaration.* More exactly, we may distinguish between the *TEI Abstract Model*, which defines a set of related concepts, and the *TEI schema* which defines a set of syntactic rules and constraints. Many (though not all) of the semantic recommendations are provided solely as informal descriptive prose, though some of them are also enforced by means of such constructs as datatypes (see *1.4.2. Datatype Specifications*), or by schema constraints expressed using the Schematron language. Schematron constraints provide information about special kinds of validation errors dependent on conditional relationships, such as the invalidity of an attribute name or value on one element when another is present, or when, given a pair of attributes such as *minOccurs* and *maxOccurs*, the minimum value exceeds the maximum value. Schematron constraints also provide warnings when a particular element, attribute, or combination is to be deprecated in future releases of the TEI scheme, to alert users that the feature in question is no longer to be used after a certain date. Very rarely, TEI Schematron constraints offer warnings that do not explicitly break with the TEI but suggest a more efficient or less ambiguous encoding. Although the descriptions, validation errors, and warnings have been written with care, there will inevitably be cases where the intention of the contributors has not been conveyed with sufficient clarity to prevent users of these Guidelines from 'extending' them in the sense of attaching slightly variant semantics to them.

Beyond this unintentional semantic extension, some of the elements described can intentionally be used in a variety of ways; for example, the element <note> has an attribute *type* which can take on arbitrary string values, depending on how it is used in a document. A new type of 'note', therefore, requires no change in the existing model. On the other hand, for many applications, it may be desirable to constrain the possible values for the *type* attribute to a small set of possibilities. A schema modified in this way would no longer necessarily regard as valid the same set of documents as the corresponding unmodified TEI schema, but would remain faithful to the same conceptual model.

This section explains how the TEI schema can be customized by suppressing elements, modifying classes of elements, or adding elements. Documents which validate against an application of the TEI schema which has been customized in this way may or may not be considered 'TEI-conformant', as further discussed in section *24.4. Conformance*.

The TEI system is designed to support modification and customization in a documented way that can be validated by an XML processor. This is achieved by writing a small TEI-conformant document, known informally as an ODD (One Document Does it all), from which an appropriate processor can generate both human-readable documentation, and a schema expressed in a language such as RELAX NG or DTD. The mechanisms used to instantiate a TEI schema differ for different schema languages, and are therefore not defined here. In XML DTDs, for example, extensive use is made of parameter entities, while in RELAX NG schemas, extensive use is made of patterns. In either case, the names of elements and, wherever possible, their attributes and content models are defined indirectly. The syntax used to implement this indirection also varies with the schema language used, but the underlying constructs in the TEI Abstract Model are given the same names. This indirection makes it simpler to implement customization of the TEI system in an accessible and human-readable way.

As further discussed in section *1. The TEI Infrastructure*, the TEI encoding schema comprises a set of class and macro declarations, and a number of *modules*. Each module is made up of element and attribute declarations, and a schema is made by combining a particular set of modules together. In the absence of any other kind of customization, when modules are combined together:

1. all the elements defined by the module (and described in the corresponding section of these Guidelines) are included in the schema;
2. each such element is identified by the canonical name given it in these Guidelines;
3. the content model of each such element is as defined by these Guidelines;
4. the names, datatypes, and permitted values declared for each attribute associated with each such element are as given in these Guidelines;
5. the elements comprising element classes and the meaning of macro declarations expressed in terms of element classes is determined by the particular combination of modules selected.

The TEI customization mechanisms allow the user to control this behaviour as follows:

1. particular elements may be suppressed, removing them from any classes in which they are members, and also from any generated schema;
2. new elements may be added to an existing class, thus making them available in macros or content models defined in terms of those classes;
3. additional attributes, or attribute values, may be specified for an individual element or for classes of elements;
4. within certain limits, attributes, or attribute values, may also be removed either from an individual element or for classes of elements;
5. the characteristics inherited by one class from another class may be modified by modifying its class membership: all members of the class then inherit the changed characteristics;
6. the set of values legal for an attribute or attribute class may be constrained or relaxed by supplying or modifying a value list, or by modifying its datatype.
7. within certain limits, the name (generic identifier) associated with an element may be changed, without changing the semantic or syntactic properties of the element;

The modification mechanisms presented in this section are quite general, and may be used to make all the types of changes just listed.

The recommended way of implementing and documenting all such modifications is by means of the ODD system described in chapter *23. Documentation Elements*; in the remainder of this section we give specific examples to illustrate how that system may be applied. An ODD processor, such as the Roma application supported by the TEI, or any other comparable set of stylesheets will use the declarations provided by an ODD to generate appropriate sets of declarations in a specific schema language such as RELAX NG or the XML DTD language. We do not discuss in detail here how this should be done, since the details are schema language-specific; some background information about the methods used for XML DTD and RELAX NG schema generation is however provided in section *1.2. Defining a TEI Schema*. Several example ODD files are also provided as part of the standard TEI release: see further section *24.3.4. Examples of Modification* below.

## 24.3.1 Kinds of Modification

For ease of discussion, we distinguish the following different kinds of modification:

1. deletion of elements;
2. modification of content models;
3. modification of attribute and attribute-value lists;
4. modification of class membership;
5. addition of new elements.
6. renaming of elements;

Each of these is described in the following sections.

Each kind of modification changes the set of documents that will be considered valid according to the resulting schema. A schema derived from any combination of unmodified TEI declarations (an "unmodified schema") may be thought of as defining a certain set of documents. A schema deriving from a combination of modified TEI declarations (a "modified schema") will define a different set of documents. The set of documents valid according to the modified schema may or may not be properly contained by the set of documents considered to be valid according to the unmodified schema. The schema `TEI-All` is the special case of the unmodified schema which includes all currently available TEI modules. We use the term *clean modification* for cases where the set of documents defined by a modified schema is a proper subset of the set of documents defined by TEI-All. Where this is not the case, that is, where the modified schema considers valid some documents which TEI-All does not, we use the term *unclean modification*. Despite this terminology, unclean modifications are not particularly deprecated, and their use may often be vital to the success of a project. The concept is introduced solely to distinguish the effects of different kinds of modification.

### 24.3.1.1 Deletion of Elements

The simplest way to modify the supplied modules is to suppress one or more of the supplied elements. This is simply done by setting the *mode* attribute to delete on an <elementSpec> for the element concerned.

For example, if the <note> element is not to be used in a particular application, but has been included via one of the supplied modules, then the schema specification concerned will contain a declaration like the following:

```
<elementSpec  ident="note"  mode="delete"/>
```

The *ident* attribute here supplies the canonical name of the element to be deleted and the *mode* attribute specifies what is to be done with it. There is no need to specify the module concerned, since element names are unique across all TEI modules. The full specification for a schema in which this modification is applied would thus be something like the following:

```
<schemaSpec  ident="mySchema">
  <moduleRef  key="core"/>
<!-- other modules used by this customization -->
  <elementSpec  ident="note"  mode="delete"/>
</schemaSpec>
```

In most cases, deletion is a clean modification, since most elements are optional. Documents that are valid with respect to the modified schema are also valid according to TEI-All. To say this another way, the set of documents matching the new schema is contained by the set of documents matching the original schema.

There are however some elements in the TEI schema which have mandatory children; for example, the element <listPerson> must contain at least one element from the `model.personLike` class. If that class has no members because all of its member elements have been removed, then the content model cannot be satisfied. A modification which keeps <listPerson> but removes all of its possible children would therefore be regarded as unclean. So long as at least one member of the class remains available, however, deleting other members would not have this effect, and would therefore be regarded as a clean modification.

In general, whenever the element deleted by a modification is mandatory within the content model of some other (undeleted) element, the result is an unclean modification, and may also break the TEI Abstract Model (*24.4.3. Conformance to the TEI Abstract Model*). However, the parent of a mandatory child can be safely removed if it is itself optional.

To determine whether or not an element is mandatory in a given context, the user must inspect the content model of the element concerned. In most cases, content models are expressed in terms of model classes rather than elements; hence, removing an element will generally be a clean modification, since there will generally be other members of the class available. If a class is completely depopulated by a modification, then the cleanliness of the modification will depend upon whether or not the class reference is mandatory or optional, in the same way as for an individual element.

#### 24.3.1.2 Modification of Content Models

The content model for an element in the TEI schema is defined by means of a <content> element within the <elementSpec> which specifies it. For example, the specification for the element <term> provided by these Guidelines contains a <content> element like the following:

```
<content>
  <macroRef  key="macro.phraseSeq"/>
</content>
```

This content model consists of a reference to a macro called `macro.phraseSeq`. Further examination shows that this macro in turn expands to an optional repeatable alternation of text (<textNode>) with references to three other classes (`model.gLike`, `model.phrase`, or `model.global`). For some particular application it might be preferable to insist that <term> elements should only contain plain text, excluding these other possibilities.[135] This could be achieved simply by supplying a specification for <term> like the following:

```
<elementSpec  ident="term"  module="core"
  mode="change">
  <content>
    <textNode/>
  </content>
</elementSpec>
```

This is a clean modification which does not change the meaning of a TEI element; there is therefore no need to assign the element to some other namespace than that of the TEI, though it may be considered good practice; see further *24.3.2. Modification and Namespaces* below.

A change of this kind, which simplifies the possible content of an element by reducing its model to one of its existing components, is always clean, because the set of documents matched by the resulting schema is a subset of the set of documents which would have been matched by TEI-All.

Note that content models are generally defined (as far as possible) in terms of references to model classes, rather than to explicit elements. This means that the need to modify content models is greatly reduced: if an element is deleted or modified, for example, then the deletion or modification will be available for every content model which

[135] Excluding `model.gLike` is generally inadvisable however, since without it the resulting schema has no way of referencing non-Unicode characters.

references that element via its class, as well as those which reference it explicitly. For this reason it is not (in general) good practice to replace class references by explicit element references, since this may have unintended side effects.

An unqualified reference to an element class within a content model generates a content model which is equivalent to an alternation of all the members of the class referenced. Thus, a content model which refers to the model class `model.phrase` will generate a content model in which any one of the members of that class is equally acceptable. The <classRef> element used to reference a class has an *expand* attribute which may be used to vary this behaviour, for example to require 'an optional repeatable alternation of all members of a class', 'a sequence containing no more than one of each member of the class', etc. as described further in *23.6. Class Specifications*.

Content model changes which are not simple restrictions on an existing model should be undertaken with caution. The set of documents matching the schema which results from such changes is unlikely to be contained by the set of documents matching TEI-All, and such changes are therefore regarded as unclean. When content models are changed or extended, care should be taken to respect the existing semantics of the element concerned as stated in these Guidelines. For example, the element <l> is defined as containing a line of verse. It would not therefore make sense to redefine its content model so that it could also include members of the class `model.pLike` such as <p> or <ab>. Although syntactically feasible, such a modification would not be regarded as TEI-conformant because it breaks the TEI Abstract Model.

#### 24.3.1.3 Modification of Attribute and Attribute Value Lists

The attributes applicable to a given element may be specified in two ways: they may be given explicitly, by means of an <attList> element within the corresponding <elementSpec>, or they may be inherited from an attribute class, as specified in the <classes> element. To add a new attribute to an element, the designer should therefore first check to see whether this attribute is already available from some existing attribute class. If it is, then the simplest method of adding it will be to make the element in question a member of that class, as further discussed below. If this is not possible, then a new <attDef> element must be added to the existing <attList> for the element in question.

Whichever method is adopted, the modification capabilities are the same as those available for elements. Attributes may be added or deleted from the list, using the *mode* attribute on <attDef> in the same way as on <elementSpec>. The 'content' of an attribute is defined by means of the <datatype>, <valList>, or <valDesc> elements within the <attDef> element. Any of these elements may be changed.

Suppose, for example, that we wish to add two attributes to the <eg> element (used to indicate examples in a text), *type* to characterize the example in some way, and *valid* to indicate whether the example is considered valid or not. A quick glance through the Guidelines indicates that the attribute class `att.typed` could be used to provide the *type* attribute, but there is no comparable class which will provide a *valid* attribute. The existing <eg> element in fact has no local attributes defined for it at all: we will therefore need to add not only an <attDef> element to define the new attribute, but also an <attList> to hold it.

We begin by adding the new *valid* attribute:

```
<elementSpec ident="eg" module="tagdocs"
  mode="change">
  <attList>
    <attDef ident="valid" mode="add"
      ns="http://www.example.com/ns/nonTEI">
      <desc>indicates whether or not the example is considered to be valid</desc>
      <datatype>
        <dataRef key="teidata.truthValue"/>
      </datatype>
    </attDef>
  </attList>
</elementSpec>
```

The value supplied for the *mode* attribute on the <attDef> element is add; if this attribute already existed on the element we are modifying this should generate an error, since a specification cannot have more than one attribute of the same name. If the attribute is already present, we can replace the whole of the existing declaration by supplying replace as the value for *mode*; alternatively, we can change some parts of an existing declaration only by supplying just the new parts, and setting change as the value for *mode*.

Because the new attribute is not defined by the TEI, it is good practice to specify a namespace for it on the <attDef>; see further *24.3.2. Modification and Namespaces*.

As noted above, adding the new *type* attribute involves changing this element's class membership; we therefore discuss that in the next section (*24.3.1.4. Class Modification*).

The canonical name for the new attribute is *valid*, and is supplied on the *ident* attribute of the <attDef> element. In this simple example, we supply only a description and datatype for the new attribute; the former is given by the <desc> element, and the latter by the <datatype> element. (There are of course many other pieces of information which could be supplied, as documented in *23. Documentation Elements*). The content of the <datatype> element is a <dataRef> element which references an existing TEI data specification.

It is often desirable to constrain the possible values for an attribute to a greater extent than is possible by simply supplying a TEI datatype for it. This facility is provided by the <valList> element, which can also appear as a child of the <attDef> element. Suppose for example that, rather than simply indicating whether or not the example is considered valid by means of the values true and false we wish to provide a more nuanced indication, using encoded values such as A, B, and C. A declaration like the following might be appropriate:

```
<elementSpec ident="eg" module="tagdocs"
  mode="change">
  <attList>
    <attDef ident="valid"
      ns="http://www.example.com/ns/nonTEI" mode="add">
      <desc>indicates the validity of the example by supplying one of three predefined
            codes for it.</desc>
      <datatype>
        <dataRef key="teidata.enumerated"/>
      </datatype>
      <valList type="closed">
        <valItem ident="A">
          <desc>validity is of the highest class</desc>
        </valItem>
        <valItem ident="B">
          <desc>validity is of the second highest class</desc>
        </valItem>
        <valItem ident="C">
          <desc>validity is of the lowest class</desc>
        </valItem>
      </valList>
    </attDef>
  </attList>
</elementSpec>
```

The same technique may be used to replace or extend the <valList> supplied as part of any attribute in the TEI schema.

#### 24.3.1.4 Class Modification

The concept of element classes was introduced in *1.3.2. Model Classes*; an understanding of it is fundamental to successful use of the TEI system. As noted there, we distinguish *model classes*, the members of which all have structural similarity, from *attribute classes*, the members of which simply share a set of attributes.

The part of an element specification which determines its class membership is an element called <classes>. All classes to which the element belongs must be specified within this, using a <memberOf> element for each.

To add an element to a class in which it is not already a member, all that is needed is to supply a new <memberOf> element within the <classes> element for the element concerned. For example, to add an element to the `att.typed` class, we include a declaration like the following:

```
<elementSpec ident="eg" module="tagdocs"
  mode="change" ns="http://www.example.com/ns/nonTEI">
```

```
  <classes  mode="change">
    <memberOf  key="att.typed"/>
  </classes>
</elementSpec>
```

Any existing class memberships for the element being changed are not affected because the *mode* attribute of the <classes> element is set to change (rather than its default value of replace). Consequently, in this case, the <eg> element retains its membership of the two classes (`model.common` and `model.graphicLike`) to which it already belongs.

Equally, to remove the attributes which an element inherits from its membership in some class, all that is needed is to remove the relevant <memberOf> element. For example, the element <term> defined in the core module is a member of two attribute classes, `att.typed` and `att.declaring`. It inherits the attributes *type* and *subtype* from the former, and the attribute *decls* from the latter. To remove the last of these attributes from this element, we need to remove it from that class:

```
<elementSpec  ident="term"  module="core"
  mode="change">
  <classes  mode="change">
    <memberOf  key="att.declaring"
      mode="delete"/>
  </classes>
</elementSpec>
```

If the intention is to change the class membership of an element completely, rather than simply add or remove it to or from one or more classes, the value of the *mode* attribute of <classes> can be set to replace (which is the default if no value is specified), indicating that the memberships indicated by its child <memberOf> elements are the only ones applicable. Thus the following declaration:

```
<elementSpec  ident="term"  module="core"
  mode="change"  ns="http://www.example.com/ns/nonTEI">
  <classes  mode="replace">
    <memberOf  key="att.interpLike"/>
  </classes>
</elementSpec>
```

would have the effect of removing the element <term> from both its existing attribute classes, and adding it to the `att.interpLike` class.

If however the *mode* attribute is set to change, the implication is that the memberships indicated by its child <memberOf> elements are to be combined with the existing memberships for the element.

To change or remove attributes inherited from an attribute class for all members of the class (as opposed to specific members of that class), it is also possible to modify the class specification itself. For example, the class `att.global.rendition` defines several attributes which are available for all elements, namely *rend*, *style*, and *rendition*. If we decide that we never wish to use the *rend* attribute, the simplest way of removing it is to supply a modified class specification for `att.global.rendition` as follows:

```
<classSpec  ident="att.global.rendition"
  type="atts"  mode="change">
  <attList>
    <attDef  ident="rend"  mode="delete"/>
  </attList>
</classSpec>
```

Because the *mode* attribute on the <classSpec> defining the attributes inherited through membership of this class has the value change, any of its existing identifiable components not specified in the modification above will remain unchanged. The only effect will therefore be to delete the *rend* attribute from the class, and hence from all elements which are members of the class.

The classes used in the TEI schema are further discussed in chapter *1. The TEI Infrastructure.* Note in particular that classes are themselves classified: the attributes inherited by a member of attribute class A may come to it directly from that class, or from another class of which A is itself a member. For example, the class `att.global` is itself a member of the classes `att.global.linking` and `att.global.analytic`. By default, these two classes are predefined as empty. However, if (for example) the `linking` module is included in a schema, a number of attributes (*corresp*, *sameAs*, etc.) are defined as members of the `att.global.linking` class. All elements which are members of `att.global` will then inherit these new attributes (see further section *1.3.1. Attribute Classes*). A new attribute may thus be added to the global class in two ways: either by adding it to the <attList> defined within the class specification for `att.global`; or by defining a new attribute class, and changing the class membership of the `att.global` class to reference it.

Such global changes should be undertaken with caution: in general removing existing non-mandatory attributes from a class will always be a clean modification, in the same way as removing non-mandatory elements. Adding a new attribute to a class however can be a clean modification only if the new attribute is labelled as belonging to some namespace other than the TEI.

The same mechanisms are available for modification of model classes. Care should be taken when modifying the model class membership of existing elements since model class membership is what determines the content model of most elements in the TEI schema, and a small change may have unintended consequences.

#### 24.3.1.5 Addition of New Elements

To add a completely new element into a schema involves providing a complete element specification for it. It is recommended that the <classes> element of this new declaration should include a reference to at least one TEI model class. Without such a reference, the new element would not otherwise be referenced by the content model of any other TEI element, and would therefore be inaccessible within a TEI document. It is also possible to modify the content models of one or more existing elements to refer to the new element explicitly, as discussed in *24.3.1.2. Modification of Content Models* but this will generally be less convenient.

For example, the three elements <bibl>, <biblFull>, and <biblStruct> are all defined as members of the class `model.biblLike`. To add a fourth member (say `<myBibl>`) to this class, we need to include in the <elementSpec> defining our new element a <memberOf> element which nominates the intended class:

```
<elementSpec ident="myBibl" mode="add"
  ns="http://www.example.com/ns/nonTEI">
  <classes>
    <memberOf key="model.biblLike"/>
  </classes>
<!-- other parts of the new declaration here -->
</elementSpec>
```

The other parts of this declaration will typically include a description for the new element and information about its content model, its attributes, etc., as further described in *23. Documentation Elements*.

#### 24.3.1.6 Renaming of Elements

Every element and other named markup construct in the TEI schema has a *canonical name*, usually composed of English language words or syllables: this name is supplied as the value of the *ident* attribute on the <elementSpec>, <attDef>, <classSpec>, or <macroSpec> used to define it. In some circumstances, it may be convenient for the element or attribute declaration used within a schema generated from that specification to take a different form, for example, permitting schemas to use generic identifiers from a different language, or in an abbreviated form. There may be many alternative identifiers for the same markup construct, and an ODD processor may choose which of them to use for a given purpose. Each such alternative name is supplied by means of an <altIdent> element within the specification element concerned.

For example, the following declaration converts <note> to `<nt>`:

```
<elementSpec ident="note" module="core"
  mode="change">
  <altIdent>nt</altIdent>
</elementSpec>
```

Note that the *mode* attribute on the <elementSpec> now takes the value change to indicate that those parts of the element specification not supplied are to be inherited from the standard definition. The content of the <altIdent> element is now available for use in place of the canonical *ident* value in the schema generated.

Renaming in this way is is an inherently unclean modification (because the set of documents matched by the resulting schema is not contained by the set matched by TEI-All), even though the process of converting any document in which elements have been renamed into an exactly equivalent document using canonical names is completely deterministic. Documents using such a schema should not claim to be in the TEI namespace, but either use a null namespace or some user-defined namespace, as further discussed in *24.3.2. Modification and Namespaces*; this is necessary to avoid the risk of name collision between the new name and all existing TEI names as well as for reasons of conformance (see further *24.3. Customization*).

### 24.3.2 Modification and Namespaces

All the elements defined by the TEI schema are labelled as belonging to a single *namespace*, maintained by the TEI and with the URI http://www.tei-c.org/ns/1.0.[136] Only elements which are unmodified or which have undergone a clean modification may use this namespace. Note however that TEI-defined attributes are not associated with any namespace.

This implies that any other modification must either specify a different namespace or, equivalently, specify no namespace at all. The *ns* attribute is provided on elements <schemaSpec>, <elementSpec>, and <attDef> for this purpose.

Suppose, for example, that we wish to add a new attribute *topic* to the existing TEI element <p>. In the absence of namespace considerations, this would be an unclean modification, since <p> does not currently have such an attribute. The most appropriate action is to explicitly attach the new attribute to a new namespace by a declaration such as the following:

```
<elementSpec ident="p" mode="change">
  <attList>
    <attDef ident="topic" mode="add"
      ns="http://www.example.org/ns/nonTEI">
      <desc>indicates the topic of a TEI paragraph</desc>
      <datatype>
        <dataRef key="teidata.word"/>
      </datatype>
    </attDef>
  </attList>
</elementSpec>
```

Document instances using a schema derived from this ODD can now indicate clearly the status of this attribute:

```
<div
   xmlns:my="http://www.example.org/ns/nonTEI">
<!-- ... -->
  <p n="12" my:topic="rabbits">Flopsy, Mopsy, Cottontail, and Peter...</p>
</div>
```

[136] This is not strictly the case, since the element <egXML> used to represent TEI examples has its own namespace, http://www.tei-c.org/ns/Examples; this is the only exception however.

Since *topic* is explicitly labelled as belonging to something other than the TEI namespace, we regard the modification which introduced it as clean. A namespace-aware processor will be able to validate those elements in the TEI namespace against the unmodified schema.[137]

Similar considerations apply when modification is made to the content model or some other aspect of an element, or when a new element is declared. All such changes should be explicitly labelled as belonging to some non-TEI namespace or to no namespace at all.

If the *ns* attribute is supplied on a <schemaSpec> element, it identifies the namespace applicable to all components of the schema being specified. Even if such a schema includes unmodified modules from the TEI namespace, the elements contained by such modules will now be regarded as belonging to the namespace specified on the <schemaSpec>. This can be useful if it is desired simply to avoid namespace processing. For example, the following schema specification results in a schema called `noName` which has no namespace, even though it comprises declarations from the TEI `header` module:

```
<schemaSpec ns="" ident="noName">
  <moduleRef key="header"/>
<!-- ... -->
</schemaSpec>
```

In addition to the TEI canonical namespace mentioned above, the TEI may also define namespaces for approved translations of the TEI schema into other languages. The namespace for such translations is the same as that for the canonical namespace, suffixed by the appropriate ISO language identifier (*vi.1 Language Identification*). A schema specification using the Chinese translation, for example, would use the namespace `http://www.tei-c.org/ns/1.0/zh`

### 24.3.3 Documenting the Modification

The elements used to define a TEI customization (<schemaSpec>, <moduleRef>, <elementSpec>, etc.) will typically be used within a TEI document which supplies further information about the intended use of the new schema, the meaning and application of any new or modified elements within it, and so on. This document will typically conform to a schema which includes components from the module described in chapter *23. Documentation Elements*.[138]

Where the customization to be documented simply consists in a selection of TEI modules, perhaps with some deletion of unwanted elements or attributes, the documentation need not specify anything further. Even here however it may be considered worthwhile to replace some of the semantic information provided by the unmodified TEI specification. For example, the <desc> element of an unmodified TEI <elementSpec> may describe an element in terms more general than appropriate to a particular project, or the <exemplum> elements within it may not illustrate the project's actual intended usage of the element, or the <remarks> element may contain discussions of matters irrelevant to the project. These elements may therefore be replaced or deleted within an <elementSpec> as necessary.

Radical revision is also possible. It is feasible to produce a modification in which the <teiHeader> or <text> elements are not required, or in which any other rule stated in these Guidelines is either not enforced or not enforceable. In fact, the mechanism, if used in an extreme way, permits replacement of all that the TEI has to say about every component of its schema. Such revisions would result in documents that are not TEI-conformant in even the broadest sense, and it is not intended that encoders use the mechanism in this way. We discuss exactly what is meant by the concept of *TEI conformance* in the next section, *24.4. Conformance*.

### 24.3.4 Examples of Modification

Several examples of customizations of the TEI are provided as part of the standard release. They include the following:

**tei_bare** The schema generated from this customization is the minimum needed for TEI Conformance. It provides only a handful of elements.

[137] Full namespace support does not exist in the DTD language, and therefore these techniques are available only to users of more modern schema languages such as RELAX NG or W3C Schema.

[138] This module can be used to document any XML schema, and has indeed been used to document several non-TEI schemas.

**tei_all** The schema generated from this customization combines all available TEI modules, providing 588 elements.

**tei_allPlus** The schema generated from this customization combines all available TEI modules with three other non-TEI vocabularies, specifically MathML, SVG, and XInclude.

It is unlikely that any project would wish to use any of these extremes unchanged. However, they form a useful starting point for customization, whether by removing modules from tei_all or tei_allPlus, or by replacing elements deleted from tei_bare. They also demonstrate how an ODD document may be constructed to provide a basic reference manual to accompany schemas generated from it.

Shortly after publication of the first edition of these Guidelines, as a demonstration of how the TEI encoding schema might be adopted to meet 90% of the needs of 90% of the TEI user community, the TEI editors produced a brief tutorial defining one specific 'clean' modification of the TEI schema, which they called TEI Lite. This tutorial and its associated DTD became very popular and are still available from the TEI web site at `http://www.tei-c.org/Guidelines/Customization/Lite/`. The tutorial and associated schema specification is also included as one of the exemplars provided with TEI P5.

An updated and expanded version of this schema known as `TEI simplePrint` was added to the Exemplars at release 3.1.0. The elements it defines have been modified to take advantage of the 'processing model' features (see further *23.5.4. Processing Models*) introduced to the Guidelines at release 3.0.0.

The exemplars provided with TEI P5 also include a customization file from which a schema for the validation of other customization files may be generated. This ODD, called tei_odds, combines the four basic modules with the tagdocs, dictionaries, gaiji, linking, and figures modules and also provides facilities for including RELAX NG or Schematron code within a document.

## 24.4 Conformance

The notion of *TEI Conformance* is intended to assist in the description of the format and contents of a particular XML document instance or set of documents. It may be found useful in such situations as:

- interchange or integration of documents amongst different researchers or users;
- software specifications for TEI-aware processing tools;
- agreements for the deposit of texts in, and distribution of texts from, archives;
- specifying the form of documents to be produced by or for a given project.

It is not intended to provide any other evaluation, for example of scholarly merit, intellectual integrity, or value for money. A document may be of major intellectual importance and yet not be TEI-conformant; a TEI-conformant document may be of no scholarly value whatsoever.

In this section we explore several aspects of conformance, and in particular attempt to define how the term *TEI-conformant* should be used. The terminology defined here should be considered normative: users and implementors of the TEI Guidelines should use the phrases 'TEI-conformant' and 'TEI Extension' only in the senses given and with the usages described.

A document is *TEI-conformant* if it:

- is a well-formed XML document (*24.4.1. Well-formedness Criterion*)
- can be validated against a *TEI Schema*, that is, a schema derived from the TEI Guidelines (*24.4.2. Validation Constraint*)
- conforms to the TEI Abstract Model (*24.4.3. Conformance to the TEI Abstract Model*)
- uses the *TEI namespace* (and other namespaces where relevant) correctly (*24.4.4. Use of the TEI Namespace*)
- is documented by means of a TEI-conformant *ODD file* (*24.4.5. Documentation Constraint*) which refers to the TEI Guidelines

Each of these criteria is discussed in more detail below.

A document is said to use a *TEI Extension* if it is a well-formed XML document which is valid against a TEI Schema which contains additional distinctions, representing concepts not present in the TEI Abstract Model, and therefore not documented in these Guidelines. Such a document cannot necessarily be transformed automatically to a TEI-conformant document without loss of information. However, since one of the goals of the TEI is to support extensions and modifications, it should not be assumed that no TEI document can include extensions: an extension

which is expressed by means of the recommended mechanisms is also a TEI-conformant document provided that those parts of it which are not extensions are TEI-conformant.

A TEI-conformant document is said to follow *TEI Recommended Practice* if, wherever these Guidelines prefer one encoding practice to another, the preferred practice is used.

### 24.4.1 Well-formedness Criterion

These Guidelines mandate the use of well-formed XML as representation format. Documents must conform to the World Wide Web Consortium recommendation of the *Extensible Markup Language (XML) 1.0 (Fourth Edition)* or successor editions found at `https://www.w3.org/TR/xml/`. Other ways of representing the concepts of the TEI Abstract Model are possible, and other representations may be considered appropriate for use in particular situations (for example, for data capture, or project-internal processing). But such alternative representations should not be considered in any way TEI-conformant.

Previous versions of these Guidelines used SGML as a representation format. With the release of P5, the only representation format supported by these Guidelines became valid XML; legacy documents in SGML format should therefore be converted using appropriate software.

A TEI-conformant document must use the TEI namespace, and therefore must also include an XML-conformant namespace declaration, as defined below (*24.4.4. Use of the TEI Namespace*).

The use of XML greatly reduces the need to consider hardware or software differences between processing environments when exchanging data. No special packing or interchange format is required for an XML document, beyond that defined by the W3C recommendations, and no special 'interchange' format is therefore proposed by these Guidelines. For discussion of encoding issues that may arise in the processing of special character sets or non-standard writing systems, see further chapter *vi Languages and Character Sets*.

In addition to the well-formedness criterion, the W3C defines the notion of a *valid* document, as being a well-formed document which matches a specific set of rules or syntactic constraints, defined by a *schema*. As noted above, TEI conformance implies that the schema used to determine validity of a given document should be derived from the present Guidelines, by means of an ODD which references and documents components which these Guidelines define.

### 24.4.2 Validation Constraint

All *TEI-conformant* documents must validate against a schema file that has been derived from the published TEI Guidelines, combined and documented in the manner described in section *24.3. Customization*. We call the formal output of this process a *TEI Schema*.

The TEI does not mandate use of any particular schema language, only that this schema[139] should have been generated from a *TEI ODD file* that references the TEI Guidelines. Currently available tools permit the expression of schemas in any or all of the XML DTD language, W3C XML Schema, and RELAX NG (both compact and XML formats). Some of what is syntactically possible using the ODD formalism cannot be represented by all schema languages; and there are some features of some schema languages which have no counterpart in ODD. No single schema language fully captures all the constraints implied by conformance to the TEI Abstract Model. A document which is valid according to a TEI schema represented using one schema language may not be valid against the same schema expressed in other languages; for example, the DTD language does not support namespaces.

As noted in section *24.3. Customization*, many varieties of TEI schema are possible and not all of them are necessarily *TEI-conformant*; derivation from an ODD is a necessary but not a sufficient condition for TEI Conformance.

### 24.4.3 Conformance to the TEI Abstract Model

The *TEI Abstract Model* is the conceptual schema instantiated by the TEI Guidelines. These Guidelines define, both formally and informally, a set of abstract concepts such as 'paragraph' or 'heading', and their structural relationships, for example stating that 'paragraph's do not contain 'heading's. These Guidelines also define classes of elements, which have both semantic and structural properties in common. Those semantic and structural properties are also a part of the TEI Abstract Model; the class membership of an existing TEI element cannot therefore be changed without changing the model. Elements can however be removed from a class by deletion, and new non-TEI elements within their own namespaces can be added to existing TEI classes.

[139] Here and elsewhere we use the word *schema* to refer to any formal document grammar language, irrespective of the formalism used to represent it.

#### 24.4.3.1 Semantic Constraints

It is an important condition of TEI conformance that elements defined in the TEI Guidelines as having one specific meaning should not be used with another. For example, the element <l> is defined in the TEI Guidelines as containing a line of verse. A schema in which it is redefined to mean a typographic line, or an ordered queue of objects of some kind, cannot therefore be TEI-conformant, whatever its other properties.

The semantics of elements defined in the TEI Guidelines are conveyed in a number of ways, ranging from formally verifiable datatypes to informal descriptive prose. In addition, a mapping between TEI elements and concepts in other conceptual models may be provided by the <equiv> element where this is available.

A schema which shares equivalent concepts to those of the TEI conceptual model may be mappable to the TEI Schema by means of such a mechanism. For example, the concept of paragraph expressed in the TEI schema by the <p> element is probably the same concept as that expressed in the DocBook schema by the `<para>` element. Such areas of overlap facilitate interchange because elements from one namespace may be readily integrated with those from another, but do not affect the definition of conformance.

A document is said to conform to the *TEI Abstract Model* if features for which an encoding is proposed by the TEI Guidelines are encoded within it using the markup and other syntactic properties defined by means of a valid *TEI-conformant* schema. Hence, even though the names of elements or attributes may vary, a TEI-conformant document must respect the TEI Semantic Model, and be valid with respect to a TEI-conformant Schema. Although it may be possible to transform a document which follows the *TEI Abstract Model* into a *TEI-conformant* document, such a document is not itself conformant.

As noted above, the notion of semantic conformance cannot be completely enforced in a formal way. The TEI conceptual model is expressed by means of formal specification in a customization file, by means of descriptive prose in the body of these Guidelines, and implicitly by examples of usage. Any inconsistency between, for example, the text of these Guidelines and a part of a specification should be considered an error and reported to the TEI Council for correction.

#### 24.4.3.2 Mandatory Components of a TEI Document

It is a long-standing requirement for any *TEI-conformant* document that it should contain a <teiHeader> element. To be more specific a *TEI-conformant* document must contain

- a single <teiHeader> element followed by one or more elements from the `model.resource` class; or
- in the case of a corpus or collection, a single overall <teiHeader> element followed by a series of <TEI> elements each with its own <teiHeader>

All <teiHeader> elements in a *TEI-conformant* document must include elements for:

**Title Statement** This should include the title of the TEI document expressed using a <titleStmt> element.

**Publication Statement** This should include the place and date of publication or distribution of the TEI document, expressed using the <publicationStmt> element.

**Source Statement** For a document derived from some previously existing document, this must include a bibliographic description of that source. For a document not so derived, this must include a brief statement that the document has no pre-existing source. In either case, this will be expressed using the <sourceDesc> element.

### 24.4.4 Use of the *TEI Namespace*

The Namespaces Recommendation of the W3C (**?**) provides a way for an XML document to combine markup from different vocabularies without risking name collision and consequent processing difficulties. While the scope of the TEI is large, there are many areas in which it makes no particular recommendation, or where it recommends that other defined markup schemas should be adopted, such as graphics or mathematics. It is also considered desirable that users of other markup schemas should be able to integrate documents using TEI markup with their own system. To meet these objectives without compromising the reliability of its encoding, a TEI-conformant document is required to make appropriate use of the TEI namespace.

Essentially all elements in a TEI Schema which represents concepts from the TEI Abstract Model belong to the TEI namespace, `http://www.tei-c.org/ns/1.0`, maintained by the TEI. A TEI-conformant document is required to declare the namespace for all the elements it contains whether these come from the TEI namespace or from other schemas.

A TEI Schema may be created which assigns TEI elements to some other namespace, or to no namespace at all. A document using such a schema cannot be considered TEI-conformant. A document which places non-TEI elements

or attributes within the TEI namespace cannot be considered TEI-conformant; such practices are strongly deprecated as they may lead to serious difficulties for processing or interchange.

### 24.4.5 Documentation Constraint

As noted in *24.4.2. Validation Constraint* above, a TEI Schema can only be generated from a TEI ODD, which also serves to document the semantics of the elements defined by it. A TEI-conformant document should therefore always be accompanied by (or refer to) a valid *TEI ODD file* specifying which modules, elements, classes, etc. are in use together with any modifications applied, and from which a TEI Schema can be generated to validate the document. The TEI supplies a number of predefined *TEI Customization exemplar ODD files* and the schemas already generated from them (see *24.3.4. Examples of Modification* ), but most projects will typically need to customize the TEI beyond what these examples provide. It is assumed, for example, that most projects will customize the TEI schema by removing those elements that are not needed for the texts they are encoding, and by providing further constraints on the attribute values and element content models the TEI provides. All such customizations must be specified by means of a valid *TEI ODD* file.

As different sorts of customization have different implications for the interchange and interoperability of TEI documents, it cannot be assumed that every customization will necessarily result in a schema that validates only TEI-conformant documents. The ODD language permits modifications which conflict with the TEI Abstract Model, even though observing this model is a requirement for TEI Conformance. The ODD language can in fact be used to describe many kinds of markup schema, including schemas which have nothing to do with the TEI at all.

Equally, it is possible to construct a TEI Schema which is identical to that derived from a given TEI ODD file without using the ODD schema. A schema can constructed simply by combining the predefined schema language fragments corresponding with the required set of TEI modules and other statements in the relevant schema language. However, the status of such a schema with respect to the `tei_all` schema cannot in general be easily determined; it may therefore be impractical to determine whether such a schema represents a clean modification or an extension. This is one reason for making the presence of a TEI ODD file a requirement for conformance.

### 24.4.6 Varieties of TEI Conformance

The conformance status of a given document may be assessed by answering the following questions, in the order indicated:

1. Is it a valid XML document, for which a TEI Schema exists? If not, then the document cannot be considered TEI-conformant in any sense.
2. Is the document accompanied by a TEI-conformant ODD specification describing its markup schema and intended semantics? If not, then the document can only be considered TEI-conformant if it validates against a predefined TEI Schema and conforms to the TEI abstract model.
3. Does the markup in the document correctly represent the TEI abstract model? Though difficult to assess, this is essential to TEI conformance.
4. Does the document claim that all of its elements come from some namespace other than the TEI (or no namespace)? If so, the document cannot be TEI-conformant.
5. If the document claims to use the TEI namespace, in part or wholly, do the elements associated with that namespace in fact belong to it? If not, the document cannot be TEI-conformant; if so, and if all non-TEI elements and attributes are correctly associated with other namespaces, then the document may be TEI-conformant.
6. Is the document valid according to a schema made by combining all TEI modules as well as valid according to the schema derived from its associated ODD specification? If so, the document is TEI-conformant.
7. Is the document valid according to the schema derived from its associated ODD specification, but not according to `tei_all`? If so, the document uses a TEI extension.
8. Is it possible automatically to transform the document into a document which is valid according to `tei_all`, using only information supplied in the accompanying ODD and without loss of information? If so, the document is TEI-conformant.

In the following table, we examine more closely some specific, though imaginary, cases:

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| Conforms to TEI Abstract Model | Y | N | Y | Y | ? | Y | N | ? |
| Valid ODD present | Y | Y | Y | Y | Y | Y | Y | N |
| Uses only non-TEI namespace(s) or none | N | N | N | N | Y | N | Y | N |
| Uses TEI and other namespaces correctly | Y | Y | N | Y | N | Y | N | Y |
| Document is valid as a subset of `tei_all` | Y | N | Y | N | N | Y | N | Y |
| Document can be converted automatically to a form which is valid as a subset of `tei_all` | Y | N | Y | N | N | Y | N | ? |

We assume firstly that each sample document assessed here is a well-formed XML document, and that it is valid against some schema.

The document in column A is TEI-conformant. Its tagging follows the TEI Abstract Model, both as regards syntactic constraints (its <l> elements appear within  elements and not the reverse) and semantic constraints (its <l> elements appear to contain verse lines rather than typographic ones). It is accompanied by a valid ODD which documents exactly how it uses the TEI. All the TEI-defined elements and attributes in the document are placed in the TEI namespace. The schema against which it is valid is a 'clean' subset of the `tei_all` schema.

The document in column B is not a TEI document. Although it is accompanied by a valid TEI ODD, the resulting schema includes some 'unclean' modifications, and represents some concepts from the TEI Abstract Model using non-TEI elements; for example, it re-defines the content model of <p> to permit  within it, and it includes an element `<pageTrimming>` which appears to have the same meaning as the existing TEI <fw> element, but the equivalence is not made explicit in the ODD. It uses the TEI namespace correctly to identify the TEI elements it contains, but the ODD does not contain enough information automatically to convert its non-TEI elements into TEI equivalents.

The document in column C is TEI conformant. It is almost the same as the document in column A, except that the names of the elements used are not those specified for the TEI namespace. Because the ODD accompanying it contains an exact mapping for each element name (using the <altIdent> element) and there are no name conflicts, it is possible to make an automatic conversion of this document.

The document in column D is a TEI Extension. It combines elements from its own namespace with unmodified TEI elements in the TEI namespace. Its usage of TEI elements conforms to the TEI Abstract Model. Its ODD defines a new `<blort>` element which has no exact TEI equivalent, but which is assigned to an existing TEI class; consequently its schema is not a clean subset of `tei_all`. If the associated ODD provided a way of mapping this element to an existing TEI element, then this would be TEI-conformant.

The document in column E is superficially similar to document D, but because it does not use any namespace declarations (or, equivalently, it assigns unmodified TEI elements to its own namespace), it may contain name collisions; there is no way of knowing whether a <p> within it is the same as the TEI's <p> or has some other meaning. The accompanying ODD file may be used to provide the human reader with information about equivalently named elements in the TEI namespace, and hence to determine whether the document is valid with respect to the TEI Abstract Model but this is not an automatable process. In particular, cases of apparent conflict (for example use of an element <p> to represent a concept not in the TEI Abstract Model but in the abstract model of some other system, whose namespace has been removed as well) cannot be reliably resolved. By our current definition therefore, this is not a TEI document.

The document in column F is TEI-conformant. The difference between it and that in column D is that the new element `<blort>` which is used in this document is a specialization of an existing TEI element, and the ODD in which it is defined specifies the mapping (a `<my:blort>` may be automatically converted to a <tei:seg type="blort">, for example). For this to work, however, the `<blort>` must observe the same syntactic constraints as the <seg>; if it does not, this would also be a case of TEI Extension.

The document in column G is not a TEI document. Its structure is fully documented by a valid TEI ODD, but it does not claim to represent the TEI Abstract Model, does not use the TEI namespace, and is not intended to validate against any TEI schema.

The document in column H is very like that in column A, but it lacks an accompanying ODD. Instead, the schema used to validate it is produced simply by combining TEI schema fragments in the same way as an ODD processor would, given the ODD. If the resulting schema is a clean subset of `tei_all`, such a document is indistinguishable

from a TEI-conformant one, but there is no way of determining (without inspection) whether this is the case if any modification or extension has been applied. Its status is therefore, like that of Text E, impossible to determine.

## 24.5 Implementation of an ODD System

This section specifies how a processing system may take advantage of the markup specification elements documented in chapter *23. Documentation Elements* of these Guidelines in order to produce project specific user documentation, schemas in one or more schema languages, and validation tools for other processors.

The specifications in this section are illustrative but not normative. Its function is to further illustrate the intended scope and application of the elements documented in chapter *23. Documentation Elements*, since it is believed that these may have application beyond the areas directly addressed by the TEI.

An ODD processing system has to accomplish two main tasks. A set of selections, deletions, changes, and additions supplied by an ODD customization (as described in *24.3. Customization*) must first be merged with the published TEI P5 ODD specifications. Next, the resulting unified ODD must be processed to produce the desired outputs.

An ODD processor is not required to do these two stages in sequence, but that may well be the simplest approach; the ODD processing tools currently provided by the TEI Consortium, which are also used to process the source of these Guidelines, adopt this approach.

### 24.5.1 Making a Unified ODD

An ODD customization must contain a single <schemaSpec> element, which defines the schema to be constructed.

`<schemaSpec>` (schema specification) generates a TEI-conformant schema and documentation for it.

- `@start` specifies entry points to the schema, i.e. which patterns may be used as the root of documents conforming to it.
- `@prefix` specifies a default prefix which will be prepended to all patterns relating to TEI elements, unless otherwise stated.
- `@targetLang` (target language) specifies which language to use when creating the objects in a schema if names for elements or attributes are available in more than one language.
- `@docLang` (documentation language) specifies which languages to use when creating documentation if the description for an element, attribute, class or macro is available in more than one language.

Amongst other attributes inherited from the `att.identified` class, this element also carries a required *ident* attribute. This provides a name for the generated schema, which other components of the processing system may use to refer to the schema being generated, e.g. in issuing error messages or as part of the generated output schema file or files. The *ns* attribute may be used to specify the default namespace within which elements valid against the resulting schema belong, as discussed in *24.3.2. Modification and Namespaces*.

The <schemaSpec> element contains an unordered series of specialized elements, each of which is of one of the following four types:

**specifications** elements from the class `model.oddDecl` (by default <elementSpec>, <classSpec>, <moduleSpec>, and <macroSpec>); these must have a *mode* attribute which determines how they will be processed.[140] If the value of *mode* is add, then the object is simply copied to the output, but if it is change, delete, or replace, then it will be looked at by other parts of the process.

**references to specifications** <specGrpRef> elements refer to <specGrp> elements that occur elsewhere in this, or another, document. A <specGrp> element, in turn, groups together a set of ODD specifications (among other things, including further <specGrpRef> elements). The use of <specGrp> and <specGrpRef> permits the ODD markup to occur at the points in documentation where they are discussed, rather than all inside <schemaSpec>. The *target* attribute of any <specGrpRef> should be followed, and the <elementSpec>, <classSpec>, and <macroSpec>, elements in the corresponding <specGrp> should be processed as described in the previous item; <specGrpRef> elements should be processed as described here.

**references to TEI Modules** <moduleRef> elements with *key* attributes refer to components of the TEI. The value of the *key* attribute matches the *ident* attribute of the <moduleSpec> element defining a TEI module. The *key*

[140] An ODD processor should recognize as erroneous such obvious inconsistencies as an attempt to include an <elementSpec> in add mode for an element which is already present in an imported module.

must be dereferenced by some means, such as reading an XML file with the TEI ODD specification (either from the local hard drive or off the Web), or looking up the reference in an XML database (again, locally or remotely); whatever means is used, it should return a stream of XML containing the element, class, and macro specifications collected together in the specified module. These specification elements are then processed in the same way as if they had been supplied directly within the <schemaSpec> being processed.

**references to external modules** a <moduleRef> element may also refer to a compatible external module by means of its *url* attribute; the content of such modules, which must be available in the RELAX NG XML syntax, are passed directly and without modification to the output schema when that is created.

Each object obtained from the TEI ODD specification using <moduleRef> by means of the *key* attribute must be checked against objects in the customization <schemaSpec> according to the following rules:

1. if there is an object in the ODD customization with the same value for the *ident* attribute, and a *mode* value of delete, then the object from the module is ignored;
2. if there is an object in the ODD customization with the same value for the *ident* attribute, and a *mode* value of replace, then the object from the module is ignored, and the one from the ODD customization is used in its place;
3. if there is an object in the ODD customization with the same value for the *ident* attribute, and a *mode* value of change, then the two objects must be merged, as described below;
4. if there is an object in the ODD customization with the same value for the *ident* attribute, and a *mode* value of add, then an error condition should be raised;
5. otherwise, the object from the module is copied to the result.

To merge two objects with the same *ident*, their component attributes and child elements must be looked at recursively. Each component may fall into one of the following four categories:

1. Some components may occur only once within the merged object (for example attributes, and <altIdent>, <content>, or <classes> elements). If such a component is found in the ODD customization, it will be copied to the output; if it is not found there, but is present in the TEI ODD specification, then that will be copied to the output.
2. Some components are grouping objects (<attList>, <valList>, for example); these are always copied to the output, and their children are then processed following the rules given in this list.
3. Some components are 'identifiable': this means that they are members of the `att.identified` class from which they inherit the *ident* attribute; examples include <attDef> and <valItem>. A component of this type will be processed according to its *mode* attribute, following the rules given above.
4. Some components may occur multiple times, but are neither grouped nor identifiable. Examples include the <desc>, <exemplum>, <remarks>, <listRef>, <datatype> and <defaultVal> elements, and the members of `model.identSynonyms`, i.e., the <equiv>, <altIdent>, and <gloss>, elements. These should be copied from both the TEI ODD specification and the ODD customization, and all occurrences included in the output.

A special problem arises with elements which are members of attribute classes, as they are permitted to override attributes inherited from a class. For example, consider this simple modification:

```
<elementSpec ident="p">
  <classes>
    <memberOf key="att.typed"/>
  </classes>
<!-- ... -->
</elementSpec>
```

The effect of its membership in the `att.typed` class is to provide <p> with a *type* attribute and a *subtype* attribute. If we wish <p> to *not* have *subtype*, we could extend the customization in our schema as follows:

```
<elementSpec ident="p">
  <classes>
    <memberOf key="att.typed"/>
  </classes>
<!-- ... -->
  <attList>
    <attDef ident="subtype" mode="delete"/>
  </attList>
</elementSpec>
```

This means that when <memberOf key="att.typed"/> is processed, that class is looked up, each attribute which it defines is examined in turn, and the customization is searched for an override. If the modification is of the attribute class itself, work proceeds as usual; if, however, the modification is at the element level, the class reference is deleted and a series of <attRef> elements is added to the element, one for each attribute inherited from the class. Since attribute classes can themselves be members of other attribute classes, membership must be followed recursively.

The effect of the concatenation of unidentifiable components should be considered carefully. An original may have

```
<elementSpec ident="p">
  <desc>marks paragraphs in prose.</desc>
<!-- ... -->
</elementSpec>
```

which would usefully be extended with this:

```
<elementSpec ident="p" mode="change">
  <desc xml:lang="es">marca párrafos en prosa.</desc>
<!-- ... -->
</elementSpec>
```

to provide an alternate description in another language. Nothing prevents the user from supplying <desc> several times in the same language, and subsequent applications will have to decide what that may mean.

Similar considerations apply to multiple example elements, though these are less likely to cause problems in documentation. Note that existing examples can only be deleted by supplying a completely new <elementSpec> in replace mode, since the <exemplum> element is not identifiable.

In the processing of the content models of elements and the content of macros, deleted elements may require special attention.[141] A content model like this:

```
<elementSpec ident="person">
<!-- ... -->
  <content>
    <alternate>
      <classRef key="model.pLike"
        minOccurs="1" maxOccurs="unbounded"/>
      <alternate minOccurs="0"
        maxOccurs="unbounded">
        <classRef key="model.personPart"/>
        <classRef key="model.global"/>
```

[141] The carthago program behind the Pizza Chef application, written by Michael Sperberg-McQueen for TEI P3 and P4, went to very great efforts to get this right. The XSLT transformations used by the P5 Roma application are not as sophisticated, partly because the RELAX NG language is more forgiving than DTDs.

```
      </alternate>
    </alternate>
  </content>
<!-- ... -->
</elementSpec>
```

requires no special treatment because everything is expressed in terms of model classes; if the class `model.personPart` is deleted explicitly, or removed because all of its members have been deleted, then `model.global` is left as the only child of the inner <alternate>. An ODD processor may or may not elect to simplify the resulting choice between nothing and `model.global` by removing the wrapper <alternate> element. However, such simplification may be considerably more complex in the general case and an ODD processor is therefore likely to be more successful in carrying out such simplification as a distinct stage during processing of ODD sources.

If an element refers directly to an element child, like this:

```
<elementSpec ident="figure">
<!-- ... -->
  <content>
    <alternate minOccurs="0"
      maxOccurs="unbounded">
      <classRef key="model.pLike"/>
      <classRef key="model.global"/>
      <elementRef key="figure"/>
      <elementRef key="figDesc"/>
      <classRef key="model.graphicLike"/>
      <classRef key="model.headLike"/>
    </alternate>
  </content>
<!-- ... -->
</elementSpec>
```

and <figDesc> has been deleted,[142] it will be necessary to remove that reference, or the resulting schema will be invalid.

The result of the work carried out should be a new <schemaSpec> which contains a complete and internally consistent set of element, class, and macro specifications, possibly also including <moduleRef> elements with *url* attributes identifying external modules.

### 24.5.2 Generating Schemas

Assuming that any modifications have been resolved, as outlined in the previous section, making a schema is now a four stage process:

1. all datatype and other macro specifications must be collected together and declared at the start of the output schema;
2. all classes must be declared in the right order (order is significant because since some classes reference others);
3. all elements are declared;
4. any <moduleRef> elements with a *url* attribute identifying an external schema must be processed.

Working in this order gives the best chance of successfully supporting all the schema languages. However, there are a number of obstacles to overcome along the way.

An ODD processor may choose to use any desired schema language or languages for its schema output, since ODD specifications are expressed as far as possible in a form independent of any schema language. The current TEI ODD processing system produces schema output in the three main schema languages as follows:

[142] Note that deletion of required elements will cause the schema specification to accept as valid documents which cannot be TEI-conformant, since they no longer conform to the TEI Abstract Model; conformance topics are addressed in more detail in *24.4. Conformance*.

- A RELAX NG (XML) schema is generated by converting content models, datatypes, and macro specifications provided within the ODD specification; a version re-expressed in the RELAX NG compact syntax is generated using James Clark's trang application.
- A DTD schema is generated by converting the RELAX NG content models to DTD language, often simplifying it to allow for the less-sophisticated output language.
- A W3C Schema schema is created by generating a RELAX NG schema and then using James Clark's trang application.

Note that the method used to generate W3C Schema means that a processor must ensure that the RELAX NG it generates follows the subset which trang is able to translate properly (see further below)—this may involve simple trial and error.

Other projects may decide to follow a different route, perhaps implementing a direct ODD to W3C Schema translator.

Secondly, it is possible to create two rather different styles of schema. On the one hand, the schema can try to maintain all the flexibility of ODD by using the facilities of the schema language for parameterization; on the other, it can remove all customization features and produce a flat result which is not suitable for further manipulation. The TEI project currently generates both styles of schema; the first as a set of schema fragments in DTD and RELAX NG languages, which can be included as modules in other schemas, and customized further; the second as the output from a processor such as Roma, in which many of the parameterization features have been removed.

The difference between the schema styles may be illustrated by considering this ODD specification:

```
<elementSpec module="drama"
  ident="performance">
<!-- ... -->
  <classes>
    <memberOf key="model.frontPart.drama"/>
  </classes>
  <content>
    <sequence>
      <alternate minOccurs="0"
        maxOccurs="unbounded">
        <classRef key="model.divTop"/>
        <classRef key="model.global"/>
      </alternate>
      <sequence minOccurs="1"
        maxOccurs="unbounded">
        <classRef key="model.common"/>
        <classRef key="model.global"
          minOccurs="0" maxOccurs="unbounded"/>
      </sequence>
      <sequence minOccurs="0"
        maxOccurs="unbounded">
        <classRef key="model.divBottom"/>
        <classRef key="model.global"
          minOccurs="0" maxOccurs="unbounded"/>
      </sequence>
    </sequence>
  </content>
<!-- ... -->
</elementSpec>
```

A simple rendering to RELAX NG produces this:

```
  performance =
element performance {
(model.divTop | model.global)*,
(model.common, model.global*)+,
(model.divBottom, model.global*)*
att.global.attribute.xmlspace,
att.global.attribute.xmlid,
att.global.attribute.n,
att.global.attribute.xmllang,
att.global.attribute.rend,
att.global.attribute.xmlbase,
att.global.linking.attribute.corresp,
att.global.linking.attribute.synch,
att.global.linking.attribute.sameAs,
att.global.linking.attribute.copyOf,
att.global.linking.attribute.next,
att.global.linking.attribute.prev,
att.global.linking.attribute.exclude,
att.global.linking.attribute.select
}
```

In the above, a subsequent redefinition of the attribute class (such as `att.global`) would have no effect, since references to such classes have been expanded to reference their constituent attributes.

The equivalent parameterized version might look like this:

```
  performance =
element performance { performance.content, performance.attributes }
performance.content =
(model.divTop | model.global)*,
(model.common, model.global*)+,
(model.divBottom, model.global*)*
performance.attributes = att.global.attributes, empty
```

Here, the attribute class `att.global` is provided via an explicit reference (`att.global.attributes`), and can therefore be redefined. Moreover, the attributes are separated from the content model, allowing either to be overridden.

In the remainder of these section, the terms *simple schema* and *parameterized schema* are used to distinguish the two schema types. An ODD processor is not required to support both, though the simple schema output is generally preferable for most applications.

Thirdly, the problem of missing components must be resolved. For example, consider this (fictitious) model for <sp>:

```
<elementSpec  ident="sp">
<!-- ... -->
  <content>
    <sequence  minOccurs="1"  maxOccurs="1">
      <classRef  key="model.global"
        minOccurs="0"  maxOccurs="unbounded"/>
      <sequence  minOccurs="0"  maxOccurs="1">
        <elementRef  key="speaker"/>
        <classRef  key="model.global"
          minOccurs="0"  maxOccurs="unbounded"/>
```

```
    </sequence>
   </sequence>
  </content>
<!-- ... -->
</elementSpec>
```

This proposes anything from the class `model.global`, followed optionally by a <speaker> element followed by anything from the `model.global` class. What happens if <speaker> is removed from the schema? The following would result:

```
<elementSpec  ident="sp">
<!-- ... -->
  <content>
   <sequence  minOccurs="1"  maxOccurs="1">
    <classRef  key="model.global"
     minOccurs="0"  maxOccurs="unbounded"/>
    <classRef  key="model.global"
     minOccurs="0"  maxOccurs="unbounded"/>
   </sequence>
  </content>
<!-- ... -->
</elementSpec>
```

which is illegal in DTD and W3C schema languages, since for a given member of `model.global` it is impossible to be sure which rule is being used. This situation is not detected when RELAX NG is used, since the language is able to cope with non-deterministic content models of this kind and does not require that only a single rule be used.

Finally, an application will need to have some method of associating the schema with document instances that use it. The TEI does not mandate any particular method of doing this, since different schema languages and processors vary considerably in their requirements. ODD processors may wish to build in support for some of the methods for associating a document instance with a schema. The TEI does not mandate any particular method, but does suggest that those which are already part of XML (the DOCTYPE declaration for DTDs) and W3C Schema (the *xsi:schemaLocation* attribute) be supported where possible.

In order for the *xsi:schemaLocation* attribute to be valid when a document is validated against either a DTD or a RELAX NG schema, ODD processors may wish to add declarations for this attribute and its namespace to the root element, even though these are not part of the TEI *per se.* For DTDs this means adding

```
  xsi:schemaLocation CDATA #IMPLIED xmlns:xsi CDATA #FIXED 'http://www.w3.org/2001/XMLSchema-inst
```

to the list of attributes on the root element, which permits the non-namespace-aware DTD language to recognize the `xsi:schemaLocation` notation. For RELAX NG, the namespace and attribute would be declared in the usual way:

```
  namespace xsi = "http://www.w3.org/2001/XMLSchema-instance"
```

and

```
attribute xsi:schemaLocation { list { teidata.namespace, teidata.pointer }+ }
```

inside the root element declaration.

Note that declaration of the *xsi:schemaLocation* attribute in a W3C Schema schema is not permitted. Therefore, if W3C Schemas are being generated by converting the RELAX NG schema (for example, with trang), it may be necessary to perform that conversion prior to adding the *xsi:schemaLocation* declaration to the RELAX NG.

It is recognized that this is an unsatisfactory solution, but it permits users to take advantage of the W3C Schema facility for indicating a schema, while still permitting documents to be validated using DTD and RELAX NG processors without any conflict.

### 24.5.3 Names and Documentation in Generated Schemas

When processing class, element, or macro specifications, there are three general rules:

1. If a RELAX NG pattern or DTD parameter entity is being created, its name is the value of the corresponding *ident* attribute, prefixed by the value of any *prefix* attribute on <schemaSpec>. This allows for elements from an external schema to be mixed in without risk of name clashes, since all TEI elements can be given a distinctive prefix such as tei_. Thus

```
<schemaSpec ident="test" prefix="tei_">
  <elementSpec ident="sp">
<!--...-->
  </elementSpec>
</schemaSpec>
```

   may generate a RELAX NG (compact syntax) pattern like this:

```
  tei_ sp = element sp { ... }
```

   References to these patterns (or, in DTDs, parameter entities) also need to be prefixed with the same value.

2. If an element or attribute is being created, its default name is the value of the *ident* attribute, but if there is an <altIdent> child, its content is used instead.

3. Where appropriate, the documentation strings in <gloss> and <desc> should be copied into the generated schema. If there is only one occurrence of either of these elements, it should be used regardless, but if there are several, local processing rules will need to be applied. For example, if there are several with different values of *xml:lang*, a locale indication in the processing environment might be used to decide which to use. For example,

```
<elementSpec module="core" ident="head">
  <equiv/>
  <gloss>heading</gloss>
  <gloss xml:lang="fr">en-tête</gloss>
  <gloss xml:lang="es">encabezamiento</gloss>
  <gloss xml:lang="it">titolo</gloss>
<!-- ... -->
</elementSpec>
```

   might generate a RELAX NG schema fragment like the following, if the locale is determined to be French:

```
      head =
    ## en-tête
    element head { head.content, head.attributes }
```

Alternatively, a selection might be made on the basis of the value of the *version* attribute which these elements carry as members of the `att.translatable` class.

In addition, there are three conventions about naming patterns relating to classes; ODD processors need not follow them, but those reading the schemas generated by the TEI project will find it necessary to understand them:

1. when a pattern for an attribute class is created, it is named after the attribute class identifier (as above) suffixed by `.attributes` (e.g. `att.editLike.attributes`);
2. when a pattern for an attribute is created, it is named after the attribute class identifier (as above) suffixed by `.attribute.` and then the identifier of the attribute (e.g. `att.editLike.attribute.resp`);
3. when a parameterized schema is created, each element generates patterns for its attributes and its contents separately, suffixing respectively `.attributes` and `.contents` to the element name.

## 24.5.4 Making a RELAX NG Schema

To create a RELAX NG schema, the processor processes every <macroSpec>, <classSpec>, and <elementSpec> in turn, creating a RELAX NG pattern for each, using the naming conventions listed above. The order of declaration is not important, and a processor may well sort them into alphabetical order of identifier.

A complete RELAX NG schema must have an `<rng:start>` element defining which elements can occur as the root of a document. The ODD <schemaSpec> has an optional *start* attribute, containing one or more patterns, which can be used to construct the `<rng:start>`. A pattern normally corresponds to an element name, but if a prefix (see above, *24.5.3. Names and Documentation in Generated Schemas*) is supplied for an element, the pattern consists of the prefix name with the element name.

### 24.5.4.1 Macros

An ODD macro generates a corresponding RELAX NG pattern by processing the body of the <content> element in the same way as elsewhere. Thus

```
<macroSpec module="tei"
  ident="macro.phraseSeq">
  <content>
    <alternate minOccurs="0"
      maxOccurs="unbounded">
      <textNode/>
      <classRef key="model.gLike"/>
      <classRef key="model.phrase"/>
      <classRef key="model.global"/>
    </alternate>
  </content>
</macroSpec>
```

produces the following

```
<rng:define name="macro.phraseSeq">
  <rng:zeroOrMore>
    <rng:choice>
      <rng:text/>
      <rng:ref name="model.gLike"/>
      <rng:ref name="model.phrase"/>
      <rng:ref name="model.global"/>
```

```
    </rng:choice>
  </rng:zeroOrMore>
</rng:define>
```

#### 24.5.4.2 Classes

An ODD model class always generates a RELAX NG pattern definition listing all the members of the class in alternation. For example

```
<classSpec module="tei" type="model"
  ident="model.measureLike">
<!-- ... -->
</classSpec>
```

will produce something like the following:

```
<rng:define name="model.measureLike">
  <rng:choice>
    <rng:ref name="num"/>
    <rng:ref name="measure"/>
    <rng:ref name="measureGrp"/>
  </rng:choice>
</rng:define>
```

assuming that the elements <num>, <measure>, and <measureGrp> are all defined in the schema concerned as members of that class. A model declaration may also generate a number of other patterns corresponding with sequences or alternations of the class members:

```
<rng:define name="model.measureLike_sequence">
  <rng:ref name="num"/>
  <rng:ref name="measure"/>
  <rng:ref name="measureGrp"/>
</rng:define>
<rng:define name="model.measureLike_sequenceOptional">
  <rng:optional>
    <rng:ref name="num"/>
  </rng:optional>
  <rng:optional>
    <rng:ref name="measure"/>
  </rng:optional>
  <rng:optional>
    <rng:ref name="measureGrp"/>
  </rng:optional>
</rng:define>
<rng:define name="model.measureLike_sequenceOptionalRepeatable">
  <rng:zeroOrMore>
    <rng:ref name="num"/>
  </rng:zeroOrMore>
  <rng:zeroOrMore>
    <rng:ref name="measure"/>
  </rng:zeroOrMore>
  <rng:zeroOrMore>
    <rng:ref name="measureGrp"/>
```

```
  </rng:zeroOrMore>
</rng:define>
<rng:define  name="model.measureLike_sequenceRepeatable">
  <rng:oneOrMore>
    <rng:ref  name="num"/>
  </rng:oneOrMore>
  <rng:oneOrMore>
    <rng:ref  name="measure"/>
  </rng:oneOrMore>
  <rng:oneOrMore>
    <rng:ref  name="measureGrp"/>
  </rng:oneOrMore>
</rng:define>
```

where the pattern name is created by appending an underscore and the name of the generation type to the class name.

When classes are referenced using the <classRef> element, it is possible to indicate which of the available patterns is required by means of the *expand* attribute.

Attribute classes work by producing a pattern containing definitions of the appropriate attributes. So

```
<classSpec  module="verse"  type="atts"
  ident="att.enjamb">
  <attList>
    <attDef  ident="enjamb"  usage="opt">
      <equiv/>
      <desc>indicates whether the end of a verse line is marked by enjambement.</desc>
      <datatype>
        <dataRef  key="teidata.enumerated"/>
      </datatype>
      <valList  type="open">
        <valItem  ident="no">
          <equiv/>
          <desc>the line is end-stopped </desc>
        </valItem>
        <valItem  ident="yes">
          <equiv/>
          <desc>the line in question runs on into the next </desc>
        </valItem>
        <valItem  ident="weak">
          <equiv/>
          <desc>the line is weakly enjambed </desc>
        </valItem>
        <valItem  ident="strong">
          <equiv/>
          <desc>the line is strongly enjambed</desc>
        </valItem>
      </valList>
    </attDef>
  </attList>
</classSpec>
```

produces

```
<rng:define  name="att.enjamb.attributes">
  <rng:ref  name="att.enjamb.attribute.enjamb"/>
  <rng:empty/>
</rng:define>
<rng:define  name="att.enjamb.attribute.enjamb">
  <rng:optional>
    <rng:attribute  name="enjamb">
      <a:documentation>(enjambement) indicates whether the end of a verse line is
marked by
           enjambement. Sample values include: 1] no; 2] yes; 3] weak; 4]
           strong</a:documentation>
      <rng:ref  name="teidata.enumerated"/>
    </rng:attribute>
  </rng:optional>
</rng:define>
```

Since the processor may have expanded the attribute classes already, separate patterns are generated for each attribute in the class as well as one for the class itself. This allows an element to refer directly to a member of a class. Notice that the <desc> element is used to add an `<a:documentation>` element to the schema, which some editors use to provide help during composition. The <desc> elements in the <valList> are used to create the human-readable sentence 'Sample values include: 1] no; 2] yes; 3] weak; 4] strong' Naturally, this behaviour is not mandatory; and other ODD processors may create documentation in other ways, or ignore those parts of the ODD specifications when creating schemas.

An individual attribute consists of an `<rng:attribute>` with a *name* attribute derived according to the naming rules described above (*24.5.3. Names and Documentation in Generated Schemas*). In addition, the ODD model supports a <defaultVal>, which is transformed to a *defaultValue* attribute in the namespace `http://relaxng.org/ns/compatibility/annotations/1.0` on the `<rng:attribute>`. The body of the attribute is taken from the <datatype> child, unless there is a supporting <valList> with a *type* value of closed. In that case an `<rng:choice>` is created, listing the allowed values. Thus the following attribute definition

```
<attDef  ident="full"  usage="opt">
  <defaultVal>yes</defaultVal>
  <valList  type="closed">
    <valItem  ident="yes">
      <desc>the name component is spelled out in full.</desc>
    </valItem>
    <valItem  ident="abb">
      <gloss>abbreviated</gloss>
      <desc>the name component is given in an abbreviated form.</desc>
    </valItem>
    <valItem  ident="init">
      <gloss>initial letter</gloss>
      <desc>the name component is indicated only by one initial.</desc>
    </valItem>
  </valList>
</attDef>
```

may generate this RELAX NG code:

```
<rng:define  name="att.full">
  <rng:optional>
    <rng:attribute  name="full"
      a:defaultValue="yes">
```

```
     <rng:choice>
       <rng:value>yes</rng:value>
       <a:documentation> the name component is spelled out in full. </a:documentation>
       <rng:value>abb</rng:value>
       <a:documentation> the name component is given in an abbreviated form.
</a:documentation>
       <rng:value>init</rng:value>
       <a:documentation> the name component is indicated only by one initial.
       </a:documentation>
     </rng:choice>
   </rng:attribute>
  </rng:optional>
</rng:define>
```

Note the use of the `http://relaxng.org/ns/compatibility/annotations/1.0` namespace to provide default values and documentation.

#### 24.5.4.3 Elements

An <elementSpec> produces a RELAX NG specification in two parts; firstly, it must generate an `<rng:define>` pattern by which other elements can refer to it, and then it must generate an `<rng:element>` with the content model and attributes. It may be convenient to make two separate patterns, one for the element's attributes and one for its content model.

The content model is created by processing the TEI declarations found within the <content> element; ; the attributes are processed in the same way as those from attribute classes, described above.

### 24.5.5 Making a DTD

Generation of DTDs largely follows the same pattern as RELAX NG generation, with one important exception—*the order of declaration matters.* A DTD may not refer to an entity which has not yet been declared. Since both macros and classes generate DTD parameter entities, the TEI Guidelines are constructed so that they can be declared in the right order. A processor must therefore work in the following order:

1. declare all model classes which have a *predeclare* value of true
2. declare all macros which have a *predeclare* value of true
3. declare all other classes
4. declare the modules (if DTD fragments are being constructed)
5. declare any remaining macros
6. declare the elements and their attributes

Let us consider a complete example, a simple element with no attributes of its own:

```
<elementSpec  module="namesdates"
  ident="faith">
  <desc>specifies the faith, religion, or belief set of a person.</desc>
  <classes>
    <memberOf  key="model.persStateLike"/>
    <memberOf  key="att.editLike"/>
    <memberOf  key="att.datable"/>
  </classes>
  <content>
    <macroRef  key="macro.phraseSeq"/>
  </content>
</elementSpec>
```

If DTD fragments are being generated (for use as described in *24.5.7. Using TEI Parameterized Schema Fragments*), this will result in the following:

```
  <!ENTITY % faith 'INCLUDE' >
<![ %faith; [

<!--doc:specifies the faith,  religion, or belief set of a person. -->
<!ELEMENT %n.faith; %om.RR; %macro.phraseSeq;>
<!ATTLIST %n.faith; xmlns CDATA "http://www.tei-c.org/ns/1.0">
<!ATTLIST %n.faith;
%att.global.attributes;
%att.editLike.attributes;
%att.datable.attributes; >
]]>
```

Here the whole stanza is contained in a marked section (for use as described in *24.5.7.2. Inclusion and Exclusion of Elements*), the element name is parameterized (see *24.5.7.3. Changing the Names of Generic Identifiers*), and the class attributes are entity references derived from the <memberOf> records in <classes>. Note the additional attribute which provides a default *xmlns* declaration for the element; the effect of this is that if the document is processed by a DTD-aware XML processor, the namespace declaration will be present automatically without the document author even being aware of it.

A simpler rendition for a flattened DTD generated from a customization will result in the following, with no containing marked section, and no parameterized name:

```
  <!ELEMENT faith %macro.phraseSeq;>
<!ATTLIST faith xmlns CDATA "http://www.tei-c.org/ns/1.0">
<!ATTLIST faith
%att.global.attribute.xmlspace;
%att.global.attribute.xmlid;
%att.global.attribute.n;
%att.global.attribute.xmllang;
%att.global.attribute.rend;
%att.global.attribute.xmlbase;
%att.global.linking.attribute.corresp;
%att.global.linking.attribute.synch;
%att.global.linking.attribute.sameAs;
%att.global.linking.attribute.copyOf;
%att.global.linking.attribute.next;
%att.global.linking.attribute.prev;
%att.global.linking.attribute.exclude;
%att.global.linking.attribute.select;
%att.editLike.attribute.cert;
%att.editLike.attribute.resp;
%att.editLike.attribute.evidence;
%att.datable.w3c.attribute.period;
%att.datable.w3c.attribute.when;
%att.datable.w3c.attribute.notBefore;
%att.datable.w3c.attribute.notAfter;
%att.datable.w3c.attribute.from;
%att.datable.w3c.attribute.to;>
```

Here the attributes from classes have been expanded into individual entity references.

### 24.5.6 Generating Documentation

In Donald Knuth's literate programming terminology (?), the previous sections have dealt with the *tangle* process; to generate documentation, we now turn to the *weave* process.

An ODD customization may consist largely of general documentation and examples, requiring no ODD-specific processing. It will normally however also contain a <schemaSpec> element and possibly some <specGrp> fragments.

The generated documentation may be of two forms. On the one hand, we may document the customization itself, that is, only those elements (etc.) which differ in their specification from that provided by the TEI reference documentation. Alternatively, we may generate reference documentation for the complete subset of the TEI which results from applying the customization. The TEI Roma tools take the latter approach, and operate on the result of the first stage processing described in *24.5.1. Making a Unified ODD*.

Generating reference documentation for <elementSpec>, <classSpec>, and <macroSpec> elements is largely dependent on the design of the preferred output. Some applications may, for example, want to turn all names of objects into hyperlinks, show lists of class members, or present lists of attributes as tables, lists, or inline prose. Another technique implemented in these Guidelines is to show lists of potential 'parents' for each element, by tracing which other elements have them as possible members of their content models.

One model of display on a web page is shown in 36, corresponding to the <faith> element shown in section *24.5.5. Making a DTD*.

Figure 36: Example reference documentation for `<faith>`

### 24.5.7 Using TEI Parameterized Schema Fragments

The TEI parameterized DTD and RELAX NG fragments make use of parameter entities and patterns for several purposes. In this section we describe their interface for the user. In general we recommend use of ODD instead of this technique, which has been retained only for compatability reasons.

#### 24.5.7.1 Selection of Modules

Special-purpose parameter entities are used to specify which modules are to be combined into a TEI DTD. They take the form TEI.xxxxx where `xxxx` is the name of the module as given in table the table on p. 2 in *1.1. TEI Modules*. For example, the parameter entity `TEI.linking` is used to define whether or not to include the module `linking`. All such parameter entities are declared by default with the value IGNORE: to select a module, therefore, the encoder declares the appropriate parameter entities with the value INCLUDE.

For XML DTD fragments, note that some modules generate two DTD fragments: for example the `analysis` module generates fragments called `analysis-decl` and `analysis`. This is because the declarations they contain are needed at different points in the creation of an XML DTD.

The parameter entity named for the module is used as the keyword controlling a conditional marked section in the DTD fragment generated by the `tei` module. The declarations for each DTD fragment constituting the module are contained within such marked sections. For example, the parameter entity `TEI.linking` appears twice in `tei.dtd`, once for the `linking-decl` schema fragment:

```
  <!ENTITY % TEI.linking 'IGNORE' >
<![%TEI.linking;[
<!ENTITY % file.linking-decl PUBLIC '-//TEI P5//ENTITIES Linking, Segmentation, and
Alignment//EN' 'linking-decl.dtd' >
%file.linking-decl;
]] >
```

and once for the `linking` schema fragment:

```
  <![%TEI.linking;[
<!ENTITY % file.linking PUBLIC '-//TEI P5//ELEMENTS Linking, Segmentation, and
Alignment//EN' 'linking.dtd' >
%file.linking;]] >
```

If TEI.linking has its default value of IGNORE, neither declaration has any effect. If however it has the value INCLUDE, then the content of each marked section is acted upon: the parameter entities `file.linking` and `file.linking-decl` are referenced, which has the effect of embedding the content of the files they represent at the appropriate point in the DTD.

The RELAX NG schema fragments can be combined in a wrapper schema using the standard mechanism of `<rng:include>` in that language.

#### 24.5.7.2 Inclusion and Exclusion of Elements

The TEI DTD fragments also use marked sections and parameter entity references to allow users to exclude the definitions of individual elements, in order either to make the elements illegal in a document or to allow the element to be redefined. The parameter entities used for this purpose have exactly the same name as the generic identifier of the element concerned. The default definition for these parameter entities is INCLUDE but they may be changed to IGNORE in order to exclude the standard element and attribute definition list declarations from the DTD.

The declarations for the element <p>, for example, are preceded by a definition for a parameter entity with the name `p` and contained within a marked section whose keyword is given as `%p;`:

```
  <!ENTITY % p 'INCLUDE' >
<![ %p; [
<!-- element and attribute list declaration for p here -->
]]
```

These parameter entities are defined immediately preceding the element whose declarations they control; because their names are completely regular, they are not documented further.

To define a DTD in which the element <p> is excluded therefore, the entity `p` needs to be redefined as IGNORE by ensuring that a declaration such as

```
  <!ENTITY % p 'IGNORE' >
```

is added earlier in the DTD than the default (see further *24.5.7.4. Embedding Local Modifications (DTD only)*).

Similarly, in the parameterized RELAX NG schemas, every element is defined by a pattern named after the element. To undefine an element therefore all that is necessary is to add a declaration like the following:

```
p = notAllowed
```

#### 24.5.7.3 Changing the Names of Generic Identifiers

In the TEI DTD fragments, elements are not referred to directly by their generic identifiers; instead, the DTD fragments refer to parameter entities which expand to the standard generic identifiers. This allows users to rename elements by redefining the appropriate parameter entity. Parameter entities used for this purpose are formed by taking the standard generic identifier of the element and attaching the string n. as a prefix. Thus the standard generic identifiers for paragraphs, notes, and quotations, <p>, <note>, and <persName> are defined by declarations of the following form:

```
<!ENTITY % n.p "p">
<!ENTITY % n.note "note">
<!ENTITY % n.persName "persName">
```

Note that since all names are case-sensitive, the specific mix of uppercase and lowercase letters in the standard generic identifier must be preserved in the entity name.

These declarations are generated by an ODD processor when TEI DTD fragments are created.

In the RELAX NG schemas, all elements are normally defined using a pattern with the same name as the element (as described in *24.5.3. Names and Documentation in Generated Schemas*): for example

```
abbr = element abbr { abbr.content, abbr.attributes }
```

The easiest way of renaming the element is thus simply to rewrite the pattern with a different element name; any references use the pattern, not the element, name.

```
abbr = element abbrev { abbr.content, abbr.attributes }
```

More complex revisions, such as redefining the content of the element (defined by the pattern `abbr.content`) or its attributes (defined by the pattern `abbr.attributes`) can be accomplished in a similar way, using the features of the RELAX NG language. The recommended method of carrying out such modifications is however to use the ODD language as further described in section *23. Documentation Elements*.

#### 24.5.7.4 Embedding Local Modifications (DTD only)

Any local modifications to a DTD (i.e. changes to a schema other than simple inclusion or exclusion of modules) are made by declarations stored in one of two local extension files, one containing modifications to the TEI parameter entities, and the other new or changed declarations of elements and their attributes. Entity declarations must be made which associate the names of these two files with the appropriate parameter entity so that the declarations they contain can be embedded within the TEI DTD at an appropriate point.

The following entities are referred to by the main `tei.dtd` file to embed portions of the TEI DTD fragments or locally developed extensions.

**`TEI.extensions.ent`** identifies a local file containing extensions to the TEI parameter entities

**`TEI.extensions.dtd`** identifies a local file containing extensions to the TEI module

For example, if the relevant files are called `project.ent` and `project.dtd`, then declarations like the following would be appropriate:

```
  <!ENTITY % TEI.extensions.ent SYSTEM 'project.ent' >
<!ENTITY % TEI.extensions.dtd SYSTEM 'project.dtd' >
```

When an entity is declared more than once, the first declaration is binding and the others are ignored. The local modifications to parameter entities should therefore be handled before the standard parameter entities themselves are declared in `tei.dtd`. The entity `TEI.extensions.ent` is referred to before any TEI declarations are handled, to allow the user's declarations to take priority. If the user does not provide a `TEI.extensions.ent` entity, the entity will be expanded to the empty string.

For example the encoder might wish to add two phrase-level elements `<it>` and `<bd>`, perhaps as synonyms for <hi rend='italics'> and <hi rend='bold'>. As described in chapter *24.3. Customization*, this involves two distinct steps: one to define the new elements, and the other to ensure that they are placed into the TEI document structure at the right place.

Creating the new declarations is done in the same way for user-defined elements as for any other; the same parameter entities need to be defined so that they may be referenced by other elements. The content models of these new elements may also reference other parameter entities, which is why they need to be declared after other declarations.

The second step involves modifying the element class to which the new elements should be attached. This requires that the parameter entity `macro.phraseSeq` should be modified to include the generic identifiers for the new elements we wish to create. The declaration for each modifiable parameter entity in the DTD includes a reference to an additional parameter entity with the same name prefixed by an `x.`; these entities are declared by default as the null string. However, in the file containing local declarations they may be redeclared to include references to the new class members:

```
  <!ENTITY % x.macro.phraseSeq 'it | bd |'>
```

and this declaration will take precedence over the default when the declaration for macro.phraseSeq is evaluated.

# A Model Classes

## A.1 About the Model Classes Appendix

This appendix gives you a list of model classes and links to the reference pages for them. There are 127 distinctly-named model classes in revision ebf0834e1 of TEI P5 Version 4.10.0 of the TEI Guidelines.

---

**model.addressLike** groups elements used to represent a postal or email address. [*1. The TEI Infrastructure*]

*Module* tei — The TEI Infrastructure

*Used by* location model.correspActionPart model.pPart.data

*Members* address affiliation email

---

**model.addrPart** groups elements such as names or postal codes which may appear as part of a postal address. [*3.6.2. Addresses*]

*Module* tei — The TEI Infrastructure

*Used by* address

*Members* model.nameLike[model.nameLike.agent[name orgName persName] model.offsetLike[geogFeat offset] model.persNamePart[addName forename genName nameLink persPronouns roleName surname] model.placeStateLike[model.placeNamePart[bloc country district geogName placeName region settlement] climate location population state terrain trait] eventName idno lang objectName rs] addrLine postBox postCode street

---

**model.annotationLike** groups elements used to represent annotations. [*17.10. The standOff Container*]

*Module* tei — The TEI Infrastructure

*Used by* listAnnotation model.standOffPart

*Members* annotation annotationBlock note

---

**model.annotationPart.body** groups elements which may be used as an <annotation> body.

*Module* tei — The TEI Infrastructure

*Used by* annotation

*Members* note ptr ref

---

**model.applicationLike** groups elements used to record application-specific information about a document in its header.

*Module* tei — The TEI Infrastructure

*Used by* appInfo

*Members* application

---

**model.attributable** groups elements that contain a word or phrase that can be attributed to a source. [*3.3.3. Quotation 4.3.2. Floating Texts*]

*Module* tei — The TEI Infrastructure

*Used by* cit macro.phraseSeq model.inter sp

*Members* model.quoteLike[cit quote] floatingText said

---

**model.availabilityPart** groups elements such as licences and paragraphs of text which may appear as part of an availability statement. [*2.2.4. Publication, Distribution, Licensing, etc.*]

*Module* tei — The TEI Infrastructure

*Used by* availability

*Members* licence

---

**model.biblLike** groups elements containing a bibliographic description. [*3.12. Bibliographic Citations and References*]

*Module* tei — The TEI Infrastructure

*Used by* broadcast cit climate event listBibl location model.entryPart.top model.inter model.msItemPart model.personPart model.standOffPart object org place population relatedItem scriptStmt sourceDesc state taxonomy terrain trait

*Members* bibl biblFull biblStruct listBibl msDesc

---

**model.biblPart** groups elements which represent components of a bibliographic description. [*3.12. Bibliographic Citations and References*]

*Module* tei — The TEI Infrastructure

*Used by* bibl

*Members* model.imprintPart[biblScope distributor pubPlace publisher] model.respLike[author editor funder meeting principal respStmt sponsor] availability bibl citedRange decoNote edition explicit extent incipit listRelation msIdentifier objectIdentifier quote relatedItem series textLang

---

**model.castItemPart** groups component elements of an entry in a cast list, such as dramatic role or actor's name.

*Module* tei — The TEI Infrastructure

*Used by* castItem

*Members* actor role roleDesc

---

**model.catDescPart** groups component elements of the TEI header Category Description.

*Module* tei — The TEI Infrastructure

*Used by* catDesc

*Members* textDesc

---

**model.certLike** groups elements which are used to indicate uncertainty or precision of other elements.

*Module* tei — The TEI Infrastructure

*Used by* certainty gap interp join precision respons space substJoin

*Members* certainty precision respons

---

**model.choicePart** groups elements (other than <choice> itself) which can be used within a <choice> alternation. [*3.5. Simple Editorial Changes*]

*Module* tei — The TEI Infrastructure

*Used by* choice

*Members* abbr am corr ex expan orig reg seg sic supplied unclear

---

**model.cmc** includes the CMC-specific element <post> and makes it available on the divPart level. This allows for using (and combining) occurrences of <post>, <u>, <incident>, and <kinesic> in one and the same  in order to be able to represent the combined use of written and spoken posts, utterances, and nonverbal acts on the GUI of multimodal CMC environments.

*Module* tei — The TEI Infrastructure

*Used by* macro.specialPara.cmc model.common

*Members* post

---

**model.common** groups common chunk- and inter-level elements. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* argument body castList div div1 div2 div3 div4 div5 div6 div7 epigraph epilogue figure performance postscript prologue set

*Members* model.cmc[post] model.divPart[model.divPart.spoken[annotationBlock u] model.lLike[l] model.pLike[ab p] eTree forest graph lg listForest sp spGrp tree] model.entryLike[entry entryFree superEntry] model.inter[model.attributable[model.quoteLike[cit quote] floatingText said] model.biblLike[bibl biblFull biblStruct listBibl msDesc] model.egLike[eg egXML] model.labelLike[desc label] model.listLike[list listApp listEvent listNym listObject listOrg listPerson listPlace listRelation listWit table] model.oddDecl[classSpec constraintSpec dataSpec elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef] model.stageLike[camera caption move sound stage tech view] castList] q

*Note* This class defines the set of chunk- and inter-level elements; it is used in many content models, including those for textual divisions.

---

**model.contentPart** groups elements which may appear as part of the content element. [*23.5.1. Defining Content Models*]

*Module* tagdocs — Documentation Elements

*Used by* alternate content sequence

*Members* alternate anyElement classRef dataRef elementRef empty macroRef sequence textNode

---

**model.correspActionPart** groups elements which define the parts (usually names, dates and places) of one action related to the correspondence.

*Module* tei — The TEI Infrastructure

*Used by* correspAction

*Members* model.addressLike[address affiliation email] model.dateLike[date time] model.nameLike[model.nameLike.agent[name orgName persName] model.offsetLike[geogFeat offset] model.persNamePart[addName forename genName nameLink persPronouns roleName surname] model.placeStateLike[model.placeNamePart[bloc country district geogName placeName region

settlement] climate location population state terrain trait] eventName idno lang objectName rs] note noteGrp

---

**model.correspContextPart** groups elements which may appear as part of the correspContext element.

*Module* tei — The TEI Infrastructure

*Used by* correspContext

*Members* model.pLike[ab p] model.ptrLike[listRef ptr ref] note noteGrp

---

**model.correspDescPart** groups together metadata elements for describing correspondence.

*Module* tei — The TEI Infrastructure

*Used by* correspDesc

*Members* correspAction correspContext note noteGrp

---

**model.dateLike** groups elements containing temporal expressions. [*3.6.4. Dates and Times 14.4. Dates*]

*Module* tei — The TEI Infrastructure

*Used by* imprint model.correspActionPart model.pPart.data model.recordingPart setting

*Members* date time

---

**model.descLike** groups elements which contain a description of their function.

*Module* tei — The TEI Infrastructure

*Used by* altGrp attDef category certainty char citeStructure classSpec constraintDecl constraintSpec dataSpec elementSpec ellipsis gap glyph graphic incident interp interpGrp join joinGrp kinesic linkGrp macroSpec media model modelGrp modelSequence moduleSpec paramSpec precision respons schemaRef schemaSpec space spanGrp substJoin taxonomy valItem vocal

*Members* desc

---

**model.describedResource** groups elements which contain the content of a digital resource and its metadata; these elements may serve as the outermost or 'root' element of a TEI-conformant document. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* teiCorpus

*Members* TEI teiCorpus

---

**model.dimLike** groups elements which describe a measurement forming part of the physical dimensions of some object.

*Module* tei — The TEI Infrastructure

*Used by* dimensions

*Members* depth height width

---

**model.div1Like** groups top-level structural divisions.
*Module* tei — The TEI Infrastructure
*Used by* back body front
*Members* div1

---

**model.div2Like** groups second-level structural divisions.
*Module* tei — The TEI Infrastructure
*Used by* div1
*Members* div2

---

**model.div3Like** groups third-level structural divisions.
*Module* tei — The TEI Infrastructure
*Used by* div2
*Members* div3

---

**model.div4Like** groups fourth-level structural divisions.
*Module* tei — The TEI Infrastructure
*Used by* div3
*Members* div4

---

**model.div5Like** groups fifth-level structural divisions.
*Module* tei — The TEI Infrastructure
*Used by* div4
*Members* div5

---

**model.div6Like** groups sixth-level structural divisions.
*Module* tei — The TEI Infrastructure
*Used by* div5
*Members* div6

---

**model.div7Like** groups seventh-level structural divisions.
*Module* tei — The TEI Infrastructure
*Used by* div6
*Members* div7

---

**model.divBottom** groups elements appearing at the end of a text division. [*4.2. Elements Common to All Divisions*]
*Module* tei — The TEI Infrastructure

*Used by* body div div1 div2 div3 div4 div5 div6 div7 epilogue figure front group lg list performance post prologue table

*Members* model.divBottomPart[closer postscript signed trailer] model.divWrapper[argument byline dateline docAuthor docDate epigraph meeting salute]

---

**model.divBottomPart** groups elements which can occur only at the end of a text division. [*4.6. Title Pages*]

*Module* tei — The TEI Infrastructure

*Used by* back model.divBottom postscript

*Members* closer postscript signed trailer

---

**model.divGenLike** groups elements used to represent a structural division which is generated rather than explicitly present in the source.

*Module* tei — The TEI Infrastructure

*Used by* body div div1 div2 div3 div4 div5 div6

*Members* divGen

---

**model.divLike** groups elements used to represent un-numbered generic structural divisions.

*Module* tei — The TEI Infrastructure

*Used by* back body div front lem rdg

*Members* div

---

**model.divPart** groups paragraph-level elements appearing directly within divisions. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* lem macro.specialPara macro.specialPara.cmc model.common rdg specGrp

*Members* model.divPart.spoken[annotationBlock u] model.lLike[l] model.pLike[ab p] eTree forest graph lg listForest sp spGrp tree

*Note* Note that this element class does not include members of the `model.inter` class, which can appear either within or between paragraph-level items.

---

**model.divPart.spoken** groups elements structurally analogous to paragraphs within spoken texts. [*8.1. General Considerations and Overview*]

*Module* spoken — Transcriptions of Speech

*Used by* model.divPart

*Members* annotationBlock u

*Note* Spoken texts may be structured in many ways; elements in this class are typically larger units such as turns or utterances.

---

**model.divTop** groups elements appearing at the beginning of a text division. [*4.2. Elements Common to All Divisions*]

*Module* tei — The TEI Infrastructure

*Used by* body castList div div1 div2 div3 div4 div5 div6 div7 epilogue group lg list performance prologue

*Members* model.divTopPart[model.headLike[head] opener signed] model.divWrapper[argument byline dateline docAuthor docDate epigraph meeting salute]

---

**model.divTopPart** groups elements which can occur only at the beginning of a text division. [*4.6. Title Pages*]

*Module* tei — The TEI Infrastructure

*Used by* model.divTop postscript

*Members* model.headLike[head] opener signed

---

**model.divWrapper** groups elements which can appear at either top or bottom of a textual division. [*4.2. Elements Common to All Divisions*]

*Module* tei — The TEI Infrastructure

*Used by* model.divBottom model.divTop

*Members* argument byline dateline docAuthor docDate epigraph meeting salute

---

**model.editorialDeclPart** groups elements which may be used inside <editorialDecl> and appear multiple times.

*Module* tei — The TEI Infrastructure

*Used by* editorialDecl

*Members* correction hyphenation interpretation normalization punctuation quotation segmentation stdVals

---

**model.egLike** groups elements containing examples or illustrations. [*23.1.1. Phrase Level Terms*]

*Module* tei — The TEI Infrastructure

*Used by* cit model.inter

*Members* eg egXML

---

**model.emphLike** groups phrase-level elements which are typographically distinct and to which a specific function can be attributed. [*3.3. Highlighting and Quotation*]

*Module* tei — The TEI Infrastructure

*Used by* model.highlighted model.limitedPhrase

*Members* code distinct emph foreign gloss ident mentioned soCalled term title

---

**model.encodingDescPart** groups elements which may be used inside <encodingDesc> and appear multiple times.

*Module* tei — The TEI Infrastructure

*Used by* encodingDesc

*Members* appInfo charDecl classDecl constraintDecl editorialDecl fsdDecl geoDecl listPrefixDef metDecl projectDesc refsDecl samplingDecl schemaRef schemaSpec styleDefDecl tagsDecl transcriptionDesc unitDecl variantEncoding

---

**model.entryLike** groups elements structurally analogous to paragraphs within dictionaries. [*10.1. Dictionary Body and Overall Structure 1.3. The TEI Class System*]

*Module* dictionaries — Dictionaries

*Used by* model.common model.standOffPart

*Members* entry entryFree superEntry

---

**model.entryPart** groups non-morphological elements appearing within a dictionary entry. [*10.1. Dictionary Body and Overall Structure*]

*Module* tei — The TEI Infrastructure

*Used by* cit dictScrap entryFree nym

*Members* colloc def etym form gramGrp hom hyph lbl orth pos pron re sense subc superEntry syll usg xr

---

**model.entryPart.top** groups high level elements within a structured dictionary entry. [*10.2. The Structure of Dictionary Entries*]

*Module* tei — The TEI Infrastructure

*Used by* entry hom re sense

*Members* model.biblLike[bibl biblFull biblStruct listBibl msDesc] cit def dictScrap entry etym form gramGrp lbl re usg xr

*Note* Members of this class typically contain related parts of a dictionary entry which form a coherent subdivision, for example a particular sense, homonym, etc.

---

**model.eventLike** groups elements which describe events.

*Module* tei — The TEI Infrastructure

*Used by* event listEvent model.orgPart model.personPart place

*Members* event listEvent

---

**model.featureVal** groups elements which represent feature values in feature structures.

*Module* tei — The TEI Infrastructure

*Used by* f fvLib if vAlt vDefault vLabel vMerge vNot vRange

*Members* model.featureVal.complex[fs vColl vMerge vNot] model.featureVal.single[binary default numeric string symbol vAlt vLabel]

---

**model.featureVal.complex** groups elements which express complex feature values in feature structures.

*Module* tei — The TEI Infrastructure

*Used by* model.featureVal
*Members* fs vColl vMerge vNot

---

**model.featureVal.single** group elements used to represent atomic feature values in feature structures.
*Module* tei — The TEI Infrastructure
*Used by* model.featureVal vColl
*Members* binary default numeric string symbol vAlt vLabel

---

**model.formPart** groups elements allowed within a <form> element in a dictionary. [*10.3.1. Information on Written and Spoken Forms*]
*Module* dictionaries — Dictionaries
*Used by* form
*Members* model.gramPart[model.lexicalRefinement[colloc gramGrp lbl pos subc usg] model.morphLike[case gen gram iType mood number per tns]] form hyph orth pron stress syll

---

**model.frontPart** groups elements which appear at the level of divisions within front or back matter. [*7.1. Front and Back Matter* ]
*Module* tei — The TEI Infrastructure
*Used by* back front
*Members* model.frontPart.drama[castList epilogue performance prologue set] divGen listBibl schemaSpec titlePage

---

**model.frontPart.drama** groups elements which appear at the level of divisions within front or back matter of performance texts only. [*7.1. Front and Back Matter* ]
*Module* tei — The TEI Infrastructure
*Used by* model.frontPart
*Members* castList epilogue performance prologue set

---

**model.fsdDeclPart** groups elements which can occur as direct children of <fsdDecl>. [*19.11. Feature System Declaration*]
*Module* iso-fs — Feature Structures
*Used by* fsdDecl
*Members* fLib fsDecl fsdLink fvLib

---

**model.gLike** groups elements used to represent individual non-Unicode characters or glyphs.
*Module* tei — The TEI Infrastructure
*Used by* am bibl byline castItem closer date dateline dictScrap docImprint entryFree etym form gramGrp head idno interp l lem line locus m macro.phraseSeq macro.specialPara macro.specialPara.cmc macro.xtext measureGrp model.paraPart msName oRef opener origDate pRef pc post rdg re sense series time trailer w witDetail xr zone

*Members* g

---

**model.global** groups elements which may appear at any point within a TEI text. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* address argument back bibl body byline castGroup castItem castList cit closer date dateline dictScrap div div1 div2 div3 div4 div5 div6 div7 docImprint docTitle entry entryFree epigraph epilogue etym figure floatingText form front gramGrp graph group head hom imprint l lem lg line list m macro.phraseSeq macro.phraseSeq.limited macro.specialPara macro.specialPara.cmc model.paraPart msItem opener origDate performance person personGrp persona post postscript prologue rdg re sense series set sourceDoc sp spGrp surface surfaceGrp table text time titlePage trailer w witDetail xr zone

*Members* model.global.edit[addSpan app damageSpan delSpan ellipsis gap space witDetail] model.global.meta[alt altGrp certainty fLib fs fvLib index interp interpGrp join joinGrp link linkGrp listTranspose precision respons span spanGrp substJoin timeline] model.global.spoken[incident kinesic pause shift vocal writing] model.milestoneLike[anchor cb fw gb lb milestone pb] model.noteLike[note noteGrp] figure metamark notatedMusic

---

**model.global.edit** groups globally available elements which perform a specifically editorial function. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* model.global

*Members* addSpan app damageSpan delSpan ellipsis gap space witDetail

---

**model.global.meta** groups globally available elements which describe the status of other elements. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* model.global model.standOffPart

*Members* alt altGrp certainty fLib fs fvLib index interp interpGrp join joinGrp link linkGrp listTranspose precision respons span spanGrp substJoin timeline

*Note* Elements in this class are typically used to hold groups of links or of abstract interpretations, or by provide indications of certainty etc. It may find be convenient to localize all metadata elements, for example to contain them within the same divison as the elements that they relate to; or to locate them all to a division of their own. They may however appear at any point in a TEI text.

---

**model.global.spoken** groups elements which may appear globally within spoken texts. [*8.1. General Considerations and Overview*]

*Module* spoken — Transcriptions of Speech

*Used by* annotationBlock model.global

*Members* incident kinesic pause shift vocal writing

*Note* This class groups elements which can appear anywhere within transcribed speech.

---

**model.gramPart** groups elements allowed within a <gramGrp> element in a dictionary. [*10.3.2. Grammatical Information*]

*Module* dictionaries — Dictionaries

*Used by* gramGrp model.formPart

*Members* model.lexicalRefinement[colloc gramGrp lbl pos subc usg] model.morphLike[case gen gram iType mood number per tns]

---

**model.graphicLike** groups elements containing images, formulae, and similar objects. [*3.10. Graphics and Other Non-textual Components*]

*Module* tei — The TEI Infrastructure

*Used by* char cit facsimile figure formula glyph model.phrase sourceDoc surface table zone

*Members* binaryObject formula graphic media

---

**model.headLike** groups elements used to provide a title or heading at the start of a text division.

*Module* tei — The TEI Infrastructure

*Used by* argument castGroup climate divGen event figure listAnnotation listApp listBibl listEvent listNym listObject listOrg listPerson listPlace listRelation listWit model.divTopPart msDesc msFrag msPart object org place population post set spGrp state table terrain trait

*Members* head

---

**model.highlighted** groups phrase-level elements which are typographically distinct. [*3.3. Highlighting and Quotation*]

*Module* tei — The TEI Infrastructure

*Used by* bibl model.phrase

*Members* model.emphLike[code distinct emph foreign gloss ident mentioned soCalled term title] model.hiLike[hi q]

---

**model.hiLike** groups phrase-level elements which are typographically distinct but to which no specific function can be attributed. [*3.3. Highlighting and Quotation*]

*Module* tei — The TEI Infrastructure

*Used by* formula m model.highlighted model.limitedPhrase model.linePart w

*Members* hi q

---

**model.identEquiv** groups elements which provide an equivalent value to the identity of their parent element, either as a formal construct or in natural language.

*Module* tei — The TEI Infrastructure

*Used by* constraintDecl model model.identSynonyms modelGrp modelSequence moduleSpec paramSpec schemaSpec

*Members* equiv gloss

---

**model.identSynonyms** groups elements which provide an explanation, equivalency, or alternative name for a markup construct.

*Module* tei — The TEI Infrastructure

*Used by* attDef classSpec constraintSpec dataSpec elementSpec macroSpec valItem

*Members* model.identEquiv[equiv gloss] altIdent

---

**model.imprintPart** groups the bibliographic elements which occur inside imprints. [*3.12. Bibliographic Citations and References*]

*Module* tei — The TEI Infrastructure

*Used by* imprint model.biblPart

*Members* biblScope distributor pubPlace publisher

---

**model.inter** groups elements which can appear either within or between paragraph-like elements. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* dictScrap entryFree etym form gramGrp head l lem macro.limitedContent macro.specialPara macro.specialPara.cmc model.common model.paraPart post rdg trailer witness xr

*Members* model.attributable[model.quoteLike[cit quote] floatingText said] model.biblLike[bibl biblFull biblStruct listBibl msDesc] model.egLike[eg egXML] model.labelLike[desc label] model.listLike[list listApp listEvent listNym listObject listOrg listPerson listPlace listRelation listWit table] model.oddDecl[classSpec constraintSpec dataSpec elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef] model.stageLike[camera caption move sound stage tech view] castList

---

**model.labelLike** groups elements used to gloss or explain other parts of a document.

*Module* tei — The TEI Infrastructure

*Used by* application climate event lg listAnnotation location model.inter notatedMusic org place population state surface terrain trait transcriptionDesc unitDef

*Members* desc label

---

**model.lexicalRefinement** elements adding further precision to the lexico-grammatical information provided for a dictionary entry.

*Module* dictionaries — Dictionaries

*Used by* model.gramPart

*Members* colloc gramGrp lbl pos subc usg

---

**model.limitedPhrase** groups phrase-level elements excluding those elements primarily intended for transcription of existing sources. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* catDesc creation macro.limitedContent macro.phraseSeq.limited witness

*Members* model.emphLike[code distinct emph foreign gloss ident mentioned soCalled term title] model.hiLike[hi q] model.pPart.data[model.addressLike[address affiliation email] model.dateLike[date time] model.measureLike[depth dim geo height measure measureGrp num unit width] model.nameLike[model.nameLike.agent[name orgName persName] model.offsetLike[geogFeat offset] model.persNamePart[addName forename genName nameLink persPronouns roleName surname]

model.placeStateLike[model.placeNamePart[bloc country district geogName placeName region settlement] climate location population state terrain trait] eventName idno lang objectName rs]] model.pPart.editorial[abbr am choice ex expan subst] model.pPart.msdesc[catchwords dimensions heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark] model.phrase.xml[att gi tag val] model.ptrLike[listRef ptr ref]

---

**model.linePart** groups transcriptional elements which appear within lines or zones of a source-oriented transcription within a <sourceDoc> element.

*Module* tei — The TEI Infrastructure

*Used by* line zone

*Members* model.hiLike[hi q] add c choice damage del handShift line mod path pc redo restore retrace seg unclear undo w zone

---

**model.listLike** groups list-like elements. [*3.8. Lists*]

*Module* tei — The TEI Infrastructure

*Used by* abstract back model.inter model.standOffPart sourceDesc sp

*Members* list listApp listEvent listNym listObject listOrg listPerson listPlace listRelation listWit table

---

**model.lLike** groups elements representing metrical components such as verse lines.

*Module* tei — The TEI Infrastructure

*Used by* head lg model.divPart model.paraPart post sp trailer

*Members* l

---

**model.lPart** groups phrase-level elements which may appear within verse only. [*6.2. Components of the Verse Line*]

*Module* tei — The TEI Infrastructure

*Used by* model.phrase w

*Members* caesura rhyme

---

**model.measureLike** groups elements which denote a number, a quantity, a measurement, or similar piece of text that conveys some numerical meaning. [*3.6.3. Numbers and Measures*]

*Module* tei — The TEI Infrastructure

*Used by* location measureGrp model.pPart.data

*Members* depth dim geo height measure measureGrp num unit width

---

**model.milestoneLike** groups milestone-style elements used to represent reference systems. [*1.3. The TEI Class System 3.11.3. Milestone Elements*]

*Module* tei — The TEI Infrastructure

*Used by* listBibl model.global org subst

*Members* anchor cb fw gb lb milestone pb

---

**model.morphLike** groups elements which provide morphological information within a dictionary entry. [*10.3. Top-level Constituents of Entries*]

*Module* dictionaries — Dictionaries

*Used by* dictScrap entryFree model.gramPart

*Members* case gen gram iType mood number per tns

---

**model.msItemPart** groups elements which can appear within a manuscript item description.

*Module* tei — The TEI Infrastructure

*Used by* msItem

*Members* model.biblLike[bibl biblFull biblStruct listBibl msDesc] model.msQuoteLike[colophon explicit finalRubric incipit rubric title] model.quoteLike[cit quote] model.respLike[author editor funder meeting principal respStmt sponsor] decoNote filiation idno msItem msItemStruct textLang

---

**model.msQuoteLike** groups elements which represent passages such as titles quoted from a manuscript as a part of its description.

*Module* tei — The TEI Infrastructure

*Used by* model.msItemPart

*Members* colophon explicit finalRubric incipit rubric title

---

**model.nameLike** groups elements which name or refer to a person, place, or organization.

*Module* tei — The TEI Infrastructure

*Used by* model.addrPart model.correspActionPart model.pPart.data org

*Members* model.nameLike.agent[name orgName persName] model.offsetLike[geogFeat offset] model.persNamePart[addName forename genName nameLink persPronouns roleName surname] model.placeStateLike[model.placeNamePart[bloc country district geogName placeName region settlement] climate location population state terrain trait] eventName idno lang objectName rs

*Note* A superset of the naming elements that may appear in datelines, addresses, statements of responsibility, etc.

---

**model.nameLike.agent** groups elements which contain names of individuals or corporate bodies. [*3.6. Names, Numbers, Dates, Abbreviations, and Addresses*]

*Module* tei — The TEI Infrastructure

*Used by* model.nameLike respStmt setting

*Members* name orgName persName

*Note* This class is used in the content model of elements which reference names of people or organizations.

---

**model.noteLike** groups globally-available note-like elements. [*3.9. Notes, Annotation, and Indexing*]

*Module* tei — The TEI Infrastructure

*Used by* adminInfo app biblStruct char climate event glyph location metDecl model.global monogr msItemStruct notesStmt object org place population rdgGrp state terrain trait

*Members* note noteGrp

---

## model.objectLike groups elements which describe objects.

*Module* tei — The TEI Infrastructure

*Used by* event listObject

*Members* listObject object

---

## model.oddDecl groups elements which generate declarations in some markup language in ODD documents.

*Module* tei — The TEI Infrastructure

*Used by* model.inter schemaSpec specGrp

*Members* classSpec constraintSpec dataSpec elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef

---

## model.oddRef groups elements which reference declarations in some markup language in ODD documents.

*Module* tei — The TEI Infrastructure

*Used by* schemaSpec specGrp

*Members* classRef dataRef elementRef macroRef moduleRef

---

## model.offsetLike groups elements which can appear only as part of a place name. [*14.2.3. Place Names*]

*Module* tei — The TEI Infrastructure

*Used by* location model.nameLike

*Members* geogFeat offset

---

## model.orgPart groups elements which form part of the description of an organization.

*Module* tei — The TEI Infrastructure

*Used by* org

*Members* model.eventLike[event listEvent] listOrg listPerson listPlace

---

## model.orgStateLike groups elements describing changeable characteristics of an organization which have a definite duration.

*Module* tei — The TEI Infrastructure

*Used by*

*Members* state

---

## model.paraPart groups elements that may appear in paragraphs and similar elements. [*3.1. Paragraphs*]

*Module* tei — The TEI Infrastructure

*Used by* macro.abContent macro.paraContent

*Members* model.gLike[g] model.global[model.global.edit[addSpan app damageSpan delSpan ellipsis gap space witDetail] model.global.meta[alt altGrp certainty fLib fs fvLib index interp interpGrp join joinGrp link linkGrp listTranspose precision respons span spanGrp substJoin timeline] model.global.spoken[incident kinesic pause shift vocal writing] model.milestoneLike[anchor cb fw gb lb milestone pb] model.noteLike[note noteGrp] figure metamark notatedMusic] model.inter[model.attributable[model.quoteLike[cit quote] floatingText said] model.biblLike[bibl biblFull biblStruct listBibl msDesc] model.egLike[eg egXML] model.labelLike[desc label] model.listLike[list listApp listEvent listNym listObject listOrg listPerson listPlace listRelation listWit table] model.oddDecl[classSpec constraintSpec dataSpec elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef] model.stageLike[camera caption move sound stage tech view] castList] model.lLike[l] model.phrase[model.graphicLike[binaryObject formula graphic media] model.highlighted[model.emphLike[code distinct emph foreign gloss ident mentioned soCalled term title] model.hiLike[hi q]] model.lPart[caesura rhyme] model.pPart.data[model.addressLike[address affiliation email] model.dateLike[date time] model.measureLike[depth dim geo height measure measureGrp num unit width] model.nameLike[model.nameLike.agent[name orgName persName] model.offsetLike[geogFeat offset] model.persNamePart[addName forename genName nameLink persPronouns roleName surname] model.placeStateLike[model.placeNamePart[bloc country district geogName placeName region settlement] climate location population state terrain trait] eventName idno lang objectName rs]] model.pPart.edit[model.pPart.editorial[abbr am choice ex expan subst] model.pPart.transcriptional[add corr damage del handShift mod orig redo reg restore retrace secl sic supplied surplus unclear undo]] model.pPart.msdesc[catchwords dimensions heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark] model.phrase.xml[att gi tag val] model.ptrLike[listRef ptr ref] model.ptrLike.form[oRef pRef] model.segLike[c cl m pc phr s seg w] model.specDescLike[specDesc specList] ruby] lg

---

## model.persNamePart groups elements which form part of a personal name. [*14.2.1. Personal Names*]

*Module* namesdates — Names, Dates, People, and Places

*Used by* model.nameLike

*Members* addName forename genName nameLink persPronouns roleName surname

---

## model.personLike groups elements which provide information about people and their relationships.

*Module* tei — The TEI Infrastructure

*Used by* event listPerson org particDesc

*Members* org person personGrp

---

## model.personPart groups elements which form part of the description of a person. [*16.2.2. The Participant Description*]

*Module* tei — The TEI Infrastructure

*Used by* person personGrp persona

*Members* model.biblLike[bibl biblFull biblStruct listBibl msDesc] model.eventLike[event listEvent] model.persStateLike[affiliation age education faith floruit gender langKnowledge nationality occupation persName persPronouns persona residence sex socecStatus state trait] birth death idno name

---

## model.persStateLike groups elements describing changeable characteristics of a person which have a definite duration, for example occupation, residence, or name.

*Module* tei — The TEI Infrastructure

*Used by* model.personPart

*Members* affiliation age education faith floruit gender langKnowledge nationality occupation persName persPronouns persona residence sex socecStatus state trait

*Note* These characteristics of an individual are typically a consequence of their own action or that of others.

---

**model.phrase** groups elements which can occur at the level of individual words or phrases. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* byline castItem closer date dateline dictScrap docImprint entryFree etym form gramGrp head l lem macro.phraseSeq macro.specialPara macro.specialPara.cmc model.paraPart opener origDate post rdg re sense time trailer witDetail xr

*Members* model.graphicLike[binaryObject formula graphic media] model.highlighted[model.emphLike[code distinct emph foreign gloss ident mentioned soCalled term title] model.hiLike[hi q]] model.lPart[caesura rhyme] model.pPart.data[model.addressLike[address affiliation email] model.dateLike[date time] model.measureLike[depth dim geo height measure measureGrp num unit width] model.nameLike[model.nameLike.agent[name orgName persName] model.offsetLike[geogFeat offset] model.persNamePart[addName forename genName nameLink persPronouns roleName surname] model.placeStateLike[model.placeNamePart[bloc country district geogName placeName region settlement] climate location population state terrain trait] eventName idno lang objectName rs]] model.pPart.edit[model.pPart.editorial[abbr am choice ex expan subst] model.pPart.transcriptional[add corr damage del handShift mod orig redo reg restore retrace secl sic supplied surplus unclear undo]] model.pPart.msdesc[catchwords dimensions heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark] model.phrase.xml[att gi tag val] model.ptrLike[listRef ptr ref] model.ptrLike.form[oRef pRef] model.segLike[c cl m pc phr s seg w] model.specDescLike[specDesc specList] ruby

*Note* This class of elements can occur within paragraphs, list items, lines of verse, etc.

---

**model.phrase.xml** groups phrase-level elements used to encode XML constructs such as element names, attribute names, and attribute values. [*23. Documentation Elements*]

*Module* tei — The TEI Infrastructure

*Used by* model.limitedPhrase model.phrase

*Members* att gi tag val

---

**model.physDescPart** groups specialized elements forming part of the physical description of a manuscript or similar written source.

*Module* msdescription — Manuscript Description

*Used by* physDesc

*Members* accMat additions bindingDesc decoDesc handDesc musicNotation objectDesc scriptDesc sealDesc typeDesc

---

**model.placeLike** groups elements used to provide information about places and their relationships.

*Module* tei — The TEI Infrastructure

*Used by* event listPlace org place settingDesc

*Members* place

---

## model.placeNamePart groups elements which form part of a place name. [*14.2.3. Place Names*]

*Module* tei — The TEI Infrastructure

*Used by* altIdentifier location model.placeStateLike msIdentifier objectIdentifier unitDef

*Members* bloc country district geogName placeName region settlement

---

## model.placeStateLike groups elements which describe changing states of a place.

*Module* tei — The TEI Infrastructure

*Used by* model.nameLike place

*Members* model.placeNamePart[bloc country district geogName placeName region settlement] climate location population state terrain trait

---

## model.pLike groups paragraph-like elements.

*Module* tei — The TEI Infrastructure

*Used by* abstract additional application availability back binding bindingDesc broadcast cRefPattern calendar climate correction correspAction correspDesc custodialHist decoDesc editionStmt editorialDecl encodingDesc equipment event exemplum front handDesc history hyphenation interpretation langKnowledge langUsage layoutDesc listRelation metDecl model.correspContextPart model.divPart msContents msDesc msFrag msItem msItemStruct msPart normalization nym object objectDesc org particDesc person personGrp persona physDesc place population post prefixDef projectDesc publicationStmt punctuation quotation recordHist recording recordingStmt refsDecl remarks samplingDecl scriptDesc scriptStmt seal sealDesc segmentation seriesStmt setting settingDesc sourceDesc sp state stdVals styleDefDecl supportDesc terrain trait transcriptionDesc typeDesc

*Members* ab p

---

## model.pLike.front groups paragraph-like elements which can occur as direct constituents of front matter. [*4.6. Title Pages*]

*Module* tei — The TEI Infrastructure

*Used by* back front

*Members* argument byline dateline docAuthor docDate docEdition docImprint docTitle epigraph head titlePart

---

## model.pPart.data groups phrase-level elements containing names, dates, numbers, measures, and similar data. [*3.6. Names, Numbers, Dates, Abbreviations, and Addresses*]

*Module* tei — The TEI Infrastructure

*Used by* bibl model.limitedPhrase model.phrase

*Members* model.addressLike[address affiliation email] model.dateLike[date time] model.measureLike[depth dim geo height measure measureGrp num unit width] model.nameLike[model.nameLike.agent[name orgName persName] model.offsetLike[geogFeat offset] model.persNamePart[addName forename genName nameLink persPronouns roleName surname] model.placeStateLike[model.placeNamePart[bloc country district geogName placeName region settlement] climate location population state terrain trait] eventName idno lang objectName rs]

---

**model.pPart.edit** groups phrase-level elements for simple editorial correction and transcription. [*3.5. Simple Editorial Changes*]

*Module* tei — The TEI Infrastructure

*Used by* bibl model.phrase pc w

*Members* model.pPart.editorial[abbr am choice ex expan subst] model.pPart.transcriptional[add corr damage del handShift mod orig redo reg restore retrace secl sic supplied surplus unclear undo]

---

**model.pPart.editorial** groups phrase-level elements for simple editorial interventions that may be useful both in transcribing and in authoring. [*3.5. Simple Editorial Changes*]

*Module* tei — The TEI Infrastructure

*Used by* model.limitedPhrase model.pPart.edit

*Members* abbr am choice ex expan subst

---

**model.pPart.msdesc** groups phrase-level elements used in manuscript description. [*11. Manuscript Description*]

*Module* tei — The TEI Infrastructure

*Used by* model.limitedPhrase model.phrase

*Members* catchwords dimensions heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark

---

**model.pPart.transcriptional** groups phrase-level elements used for editorial transcription of pre-existing source materials. [*3.5. Simple Editorial Changes*]

*Module* tei — The TEI Infrastructure

*Used by* am lg model.pPart.edit

*Members* add corr damage del handShift mod orig redo reg restore retrace secl sic supplied surplus unclear undo

---

**model.profileDescPart** groups elements which may be used inside <profileDesc> and appear multiple times.

*Module* tei — The TEI Infrastructure

*Used by* profileDesc

*Members* abstract calendarDesc correspDesc creation handNotes langUsage listTranspose particDesc settingDesc textClass textDesc

---

**model.ptrLike** groups elements used for purposes of location and reference. [*3.7. Simple Links and Cross-References*]

*Module* tei — The TEI Infrastructure

*Used by* analytic application bibl biblStruct cit eLeaf eTree entry event listRef model.correspContextPart model.limitedPhrase model.phrase model.publicationStmtPart.detail monogr notatedMusic org person place relatedItem series transcriptionDesc

*Members* listRef ptr ref

---

**model.ptrLike.form** groups elements used for purposes of location of particular orthographic or pronunciation forms within a dictionary entry. [*10. Dictionaries*]

*Module* dictionaries — Dictionaries

*Used by* model.phrase

*Members* oRef pRef

---

**model.publicationStmtPart.agency** groups the child elements of a <publicationStmt> element of the TEI header that indicate an authorising agent. [*2.2.4. Publication, Distribution, Licensing, etc.*]

*Module* tei — The TEI Infrastructure

*Used by* publicationStmt

*Members* authority distributor publisher

*Note* The 'agency' child elements, while not required, are required if one of the 'detail' child elements is to be used. It is not valid to have a 'detail' child element without a preceding 'agency' child element.
See also `model.publicationStmtPart.detail`.

---

**model.publicationStmtPart.detail** groups the agency-specific child elements of the <publicationStmt> element of the TEI header. [*2.2.4. Publication, Distribution, Licensing, etc.*]

*Module* tei — The TEI Infrastructure

*Used by* publicationStmt

*Members* model.ptrLike[listRef ptr ref] address availability date idno pubPlace

*Note* A 'detail' child element may not occur unless an 'agency' child element precedes it.
See also `model.publicationStmtPart.agency`.

---

**model.quoteLike** groups elements used to directly contain quotations.

*Module* tei — The TEI Infrastructure

*Used by* model.attributable model.msItemPart

*Members* cit quote

---

**model.rdgLike** groups elements which contain a single reading, other than the lemma, within a textual variation. [*13.1. The Apparatus Entry, Readings, and Witnesses*]

*Module* textcrit — Critical Apparatus

*Used by* app rdgGrp

*Members* rdg

*Note* This class allows for variants of the <rdg> element to be easily created via TEI customizations.

---

**model.rdgPart** groups elements which mark the beginning or ending of a fragmentary manuscript or other witness. [*13.1.5. Fragmentary Witnesses*]

*Module* textcrit — Critical Apparatus

*Used by* lem rdg

*Members* lacunaEnd lacunaStart wit witEnd witStart

*Note* These elements may appear anywhere within the elements <lem> and <rdg>, and also within any of their constituent elements.

---

**model.recordingPart** groups elements used to describe details of an audio or video recording. [*8.2. Documenting the Source of Transcribed Speech*]

*Module* spoken — Transcriptions of Speech

*Used by* recording

*Members* model.dateLike[date time] broadcast equipment media respStmt

---

**model.resource** groups separate elements which constitute the content of a digital resource, as opposed to its metadata. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* TEI teiCorpus

*Members* facsimile fsdDecl sourceDoc standOff text

---

**model.respLike** groups elements which are used to indicate intellectual or other significant responsibility, for example within a bibliographic element.

*Module* tei — The TEI Infrastructure

*Used by* editionStmt model.biblPart model.msItemPart titleStmt

*Members* author editor funder meeting principal respStmt sponsor

---

**model.segLike** groups elements used for arbitrary segmentation. [*17.3. Blocks, Segments, and Anchors 18.1. Linguistic Segment Categories*]

*Module* tei — The TEI Infrastructure

*Used by* bibl model.phrase

*Members* c cl m pc phr s seg w

*Note* The principles on which segmentation is carried out, and any special codes or attribute values used, should be defined explicitly in the <segmentation> element of the <encodingDesc> within the associated TEI header.

---

**model.settingPart** groups elements used to describe the setting of a linguistic interaction.

*Module* tei — The TEI Infrastructure

*Used by* setting

*Members* activity locale placeName

---

**model.sourceDescPart** groups elements which may be used inside <sourceDesc> and appear multiple times.

*Module* tei — The TEI Infrastructure

*Used by* sourceDesc

*Members* recordingStmt scriptStmt

---

**model.specDescLike** groups elements for referring to specification elements. [*23. Documentation Elements*]

*Module* tei — The TEI Infrastructure

*Used by* model.phrase

*Members* specDesc specList

---

**model.stageLike** groups elements containing stage directions or similar things defined by the module for performance texts. [*7.3. Other Types of Performance Text*]

*Module* tei — The TEI Infrastructure

*Used by* lg model.inter sp spGrp

*Members* camera caption move sound stage tech view

*Note* Stage directions are members of class *inter*: that is, they can appear between or within component-level elements.

---

**model.standOffPart** groups elements which may be used as children of <standOff>.

*Module* tei — The TEI Infrastructure

*Used by* standOff

*Members* model.annotationLike[annotation annotationBlock note] model.biblLike[bibl biblFull biblStruct listBibl msDesc] model.entryLike[entry entryFree superEntry] model.global.meta[alt altGrp certainty fLib fs fvLib index interp interpGrp join joinGrp link linkGrp listTranspose precision respons span spanGrp substJoin timeline] model.listLike[list listApp listEvent listNym listObject listOrg listPerson listPlace listRelation listWit table] castList eTree forest graph listAnnotation listChange listForest seg tree u xenoData zone

---

**model.teiHeaderPart** groups high level elements which may appear more than once in a TEI header.

*Module* tei — The TEI Infrastructure

*Used by* teiHeader

*Members* encodingDesc profileDesc xenoData

---

**model.textDescPart** groups elements used to categorize a text for example in terms of its situational parameters.

*Module* tei — The TEI Infrastructure

*Used by* textDesc

*Members* channel constitution derivation domain factuality interaction preparedness

---

**model.titlepagePart** groups elements which can occur as direct constituents of a title page, such as <docTitle>, <docAuthor>, <docImprint>, or <epigraph>. [*4.6. Title Pages*]

*Module* tei — The TEI Infrastructure

*Used by* msItem titlePage

*Members* argument binaryObject byline docAuthor docDate docEdition docImprint docTitle epigraph graphic imprimatur titlePart

# B Attribute Classes

## B.1 About the Attribute Classes Appendix

This appendix gives you a list of attribute classes and links to the reference pages for them. There are 85 distinctly-named attribute classes in revision ebf0834e1 of TEI P5 Version 4.10.0 of the TEI Guidelines.

---

**att.anchoring** (anchoring) provides attributes for use on annotations, e.g. notes and groups of notes describing the existence and position of an anchor for annotations.

*Module* tei — The TEI Infrastructure

*Members* note noteGrp

*Attributes*

`@anchored` (anchored) indicates whether the copy text shows the exact place of reference for the note.

*Status* Optional

*Datatype* teidata.truthValue

*Default* true

*Note* In modern texts, notes are usually anchored by means of explicit footnote or endnote symbols. An explicit indication of the phrase or line annotated may however be used instead (e.g. 'page 218, lines 3–4'). The *anchored* attribute indicates whether any explicit location is given, whether by symbol or by prose cross-reference. The value true indicates that such an explicit location is indicated in the copy text; the value false indicates that the copy text does not indicate a specific place of attachment for the note. If the specific symbols used in the copy text at the location the note is anchored are to be recorded, use the *n* attribute.

`@targetEnd` (target end) points to the end of the span to which the note is attached, if the note is not embedded in the text at that point.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Note* This attribute is retained for backwards compatibility; it may be removed at a subsequent release of the Guidelines. The recommended way of pointing to a span of elements is by means of the `range` function of XPointer, as further described in *17.2.4.6. range()*.

*Example*

```
<p>(...) tamen reuerendos dominos archiepiscopum et canonicos Leopolienses
necnon episcopum in duplicibus Quatuortemporibus<anchor xml:id="A55234"/>
totaliter expediui...</p>
<!-- elsewhere in the document -->
<noteGrp targetEnd="#A55234">
  <note xml:lang="en"> Quatuor Tempora, so called dry fast days.
  </note>
  <note xml:lang="pl"> Quatuor Tempora, tzw. Suche dni postne.
  </note>
</noteGrp>
```

---

**att.ascribed** provides attributes for elements representing speech or action that can be ascribed to a specific individual. [*3.3.3. Quotation 8.3. Elements Unique to Spoken Texts*]

*Module* tei — The TEI Infrastructure

*Members* att.ascribed.directed[kinesic move pause q said sp spGrp stage u vocal writing] annotationBlock change incident post setting shift

*Attributes*

`@who` indicates the person, or group of people, to whom the element content is ascribed.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

In the following example from Hamlet, speeches (<sp>) in the body of the play are linked to <role> elements in the <castList> using the *who* attribute.

```
<castItem type="role">
  <role xml:id="Barnardo">Bernardo</role>
</castItem>
<castItem type="role">
  <role xml:id="Francisco">Francisco</role>
  <roleDesc>a soldier</roleDesc>
</castItem>
<!-- ... -->
<sp who="#Barnardo">
  <speaker>Bernardo</speaker>
  <l n="1">Who's there?</l>
</sp>
<sp who="#Francisco">
  <speaker>Francisco</speaker>
  <l n="2">Nay, answer me: stand, and unfold yourself.</l>
</sp>
```

*Note* For transcribed speech, this will typically identify a participant or participant group; in other contexts, it will point to any identified <person> element.

---

## att.ascribed.directed provides attributes for elements representing speech or action that can be directed at a group or individual. [*3.3.3. Quotation 8.3. Elements Unique to Spoken Texts*]

*Module* tei — The TEI Infrastructure

*Members* kinesic move pause q said sp spGrp stage u vocal writing

*Attributes*

- att.ascribed
  - *@who*

`@toWhom` indicates the person, or group of people, to whom a speech act or action is directed.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

In the following example from Mary Pix's The False Friend, speeches (<sp>) in the body of the play are linked to <castItem> elements in the <castList> using the *toWhom* attribute, which is used to specify who the speech is directed to. Additionally, the <stage> includes *toWhom* to indicate the directionality of the action.

```
<castItem type="role">
  <role xml:id="emil">Emilius.</role>
</castItem>
<castItem type="role">
  <role xml:id="lov">Lovisa</role>
</castItem>
<castItem type="role">
  <role xml:id="serv">A servant</role>
</castItem>
<!-- ... -->
<sp who="#emil"
  toWhom="#lov">
```

```
  <speaker>Emil.</speaker>
  <l  n="1">My love!</l>
</sp>
<sp  who="#lov"
  toWhom="#emil">
  <speaker>Lov.</speaker>
  <l  n="2">I have no Witness of my Noble Birth</l>
  <stage  who="emil"
    toWhom="#serv">Pointing to her Woman.</stage>
  <l>But that poor helpless wretch——</l>
</sp>
```

*Note* To indicate the recipient of written correspondence, use the elements used in section *2.4.6. Correspondence Description*, rather than a *toWhom* attribute.

---

**att.breaking** provides attributes to indicate whether or not the element concerned is considered to mark the end of an orthographic token in the same way as whitespace. [*3.11.3. Milestone Elements*]

*Module* tei — The TEI Infrastructure

*Members* cb gb lb milestone pb

*Attributes*

`@break` indicates whether or not the element bearing this attribute should be considered to mark the end of an orthographic token in the same way as whitespace.

*Status* Recommended

*Datatype* teidata.enumerated

*Sample values include* **yes** the element bearing this attribute is considered to mark the end of any adjacent orthographic token irrespective of the presence of any adjacent whitespace

**no** the element bearing this attribute is considered not to mark the end of any adjacent orthographic token irrespective of the presence of any adjacent whitespace

**maybe** the encoding does not take any position on this issue.

In the following lines from the 'Dream of the Rood', the words *lāðost* and *reord-berendum* each start on one line and continue onto the next.

```
<ab> ...epesa
tome iu icpæs ȝeporden þita heardoſt .
leodum la<lb  break="no"/> ðost ærþan ichim lifes
þeȝ rihtne ȝerymde reord be<lb  break="no"/>
rendum hpæt me þaȝepeorðode þuldres ealdor ofer...
</ab>
```

---

**att.calendarSystem** provides attributes for indicating calendar systems to which a date belongs. [*3.6.4. Dates and Times 14.4. Dates*]

*Module* tei — The TEI Infrastructure

*Members* date docDate origDate time

*Attributes*

`@calendar` indicates one or more systems or calendars to which the date represented by the content of this element belongs.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Schematron* <sch:rule context="tei:*[@calendar]"> <sch:assert test="string-length(

normalize-space(.) ) gt 0"> @calendar indicates one or more systems or calendars to which the date represented by the content of this element belongs, but this <sch:name/> element has no textual content.</sch:assert> </sch:rule>

```
He was born on <date  calendar="#gregorian">Feb. 22, 1732</date>
(<date  calendar="#julian"
  when="1732-02-22">Feb. 11, 1731/32,
O.S.</date>).

He was born on <date  calendar="#gregorian #julian"
  when="1732-02-22">Feb. 22, 1732
(Feb. 11, 1731/32, O.S.)</date>.
```

*Note* Note that the *calendar* attribute declares the calendar system used to interpret the textual content of an element, as it appears on an original source. It does *not* modify the interpretation of the normalization attributes provided by att.datable.w3c, att.datable.iso, or att.datable.custom. Attributes from those first two classes are always interpreted as Gregorian or proleptic Gregorian dates, as per the respective standards on which they are based. The calender system used to interpret the last (att.datable.custom) may be specified with *datingMethod*.

---

## att.canonical provides attributes that can be used to associate a representation such as a name or title with canonical information about the object being named or referenced. [*14.1.1. Linking Names and Their Referents*]

*Module* tei — The TEI Infrastructure

*Members* att.naming[att.personal[addName eventName forename genName name objectName orgName persName placeName roleName surname] affiliation author birth bloc climate collection country death district editor education event geogFeat geogName institution nationality occupation offset origPlace population pubPlace region repository residence rs settlement socecStatus state terrain trait] actor authority bibl biblFull biblStruct catDesc correspDesc date distributor docAuthor docTitle faith funder material meeting object objectType post principal publisher relation resp respStmt sponsor term time title unitDecl unitDef

*Attributes*

`@key` provides an externally-defined means of identifying the entity (or entities) being named, using a coded value of some kind.

*Status* Optional

*Datatype* teidata.text

```
<author>
  <name  key="Hugo, Victor (1802-1885)"
    ref="http://www.idref.fr/026927608">Victor Hugo</name>
</author>
```

*Note* The value may be a unique identifier from a database, or any other externally-defined string identifying the referent. No particular syntax is proposed for the values of the *key* attribute, since its form will depend entirely on practice within a given project.

`@ref` (reference) provides an explicit means of locating a full definition or identity for the entity being named by means of one or more URIs.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

```
<name ref="http://viaf.org/viaf/109557338"
  type="person">Seamus Heaney</name>
```

*Note* The value must point directly to one or more XML elements or other resources by means of one or more URIs, separated by whitespace. If more than one is supplied the implication is that the name identifies several distinct entities.

*Example* In this contrived example, a canonical reference to the same organisation is provided in four different ways.

```
<author n="1">
  <name ref="http://nzetc.victoria.ac.nz/tm/scholarly/name-427308.html"
    type="organisation">New Zealand Parliament, Legislative Council</name>
</author>

<author n="2">
  <name ref="nzvn:427308"
    type="organisation">New Zealand Parliament, Legislative Council</name>
</author>

<author n="3">
  <name ref="./named_entities.xml#o427308"
    type="organisation">New Zealand Parliament, Legislative Council</name>
</author>


<author n="4">
  <name key="name-427308"
    type="organisation">New Zealand Parliament, Legislative Council</name>
</author>
```

The first presumes the availability of an internet connection and a processor that can resolve a URI (most can). The second requires, in addition, a <prefixDef> that declares how the `nzvm` prefix should be interpreted. The third does not require an internet connection, but does require that a file named named_entities.xml be in the same directory as the TEI document. The fourth requires that an entire external system for key resolution be available.

*Note* The *key* attribute is more flexible and general-purpose, but its use in interchange requires that documentation about how the key is to be resolved be sent to the recipient of the TEI document. In contrast values of the *ref* attribute are resolved using the widely accepted protocols for a URI, and thus less documentation, if any, is likely required by the recipient in data interchange.

These guidelines provide no semantic basis or suggested precedence when both *key* and *ref* are provided. For this reason simultaneous use of both is not recommended unless documentation explaining the use is provided, probably in an ODD customization, for interchange.

---

## att.citeStructurePart provides attributes for selecting particular elements within a document.

*Module* header — The TEI Header

*Members* citeData citeStructure

*Attributes*

`@use` (use) supplies an XPath selection pattern using the syntax defined in ?. The XPath pattern is relative to the context given in *match*, which will either be a sibling attribute in the case of <citeStructure> or on the parent <citeStructure> in the case of <citeData>.

*Status* Required

*Datatype* teidata.xpath

**att.citing** provides attributes for specifying the specific part of a bibliographic item being cited. [*1.3.1. Attribute Classes*]

*Module* tei — The TEI Infrastructure

*Members* biblScope citedRange

*Attributes*

`@unit` identifies the unit of information conveyed by the element.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **volume** (volume) the element contains a volume number.
**issue** the element contains an issue number, or volume and issue numbers.
**page** (page) the element contains a page number or page range.
**line** the element contains a line number or line range.
**chapter** (chapter) the element contains a chapter indication (number and/or title)
**part** the element identifies a part of a book or collection.
**column** the element identifies a column.
**entry** the element identifies an entry number or label in a list of entries.

`@from` specifies the starting point of the range of units indicated by the *unit* attribute.
*Status* Optional
*Datatype* teidata.word

`@to` specifies the end-point of the range of units indicated by the *unit* attribute.
*Status* Optional
*Datatype* teidata.word

---

**att.cmc** (computer-mediated communication) provides attributes categorizing how the element content was created in a CMC environment.

*Module* tei — The TEI Infrastructure

*Members* ab abbr add addName address affiliation alt altGrp anchor argument bibl biblFull biblStruct binaryObject bloc byline c cb choice cit cl climate closer corr country date dateline del desc distinct district docAuthor docDate email emph epigraph expan fLib figure floatingText foreign forename formula fs fvLib gap gb genName geo geogFeat geogName gloss graphic head hi idno incident index interp interpGrp join joinGrp kinesic l label lb lg link linkGrp list listBibl listEvent listNym listObject listOrg listPerson listPlace listRelation location m measure measureGrp media meeting mentioned milestone name nameLink notatedMusic note noteGrp num objectName offset opener orgName orig p pause pb pc persName persPronouns phr placeName population post postscript ptr q quote ref reg region roleName rs ruby s said salute seg settlement shift sic signed soCalled span spanGrp stage state surname table term terrain time timeline title trailer trait unclear unit vocal w writing

*Attributes*

`@generatedBy` (generated by) categorizes how the content of an element was generated in a CMC environment.
*Status* Optional
*Datatype* teidata.enumerated
*Schematron* <sch:rule context="tei:*[@generatedBy]">
<sch:assert test="ancestor-or-self::tei:post">The @generatedBy attribute is for use within a <post> element.</sch:assert> </sch:rule>
*Suggested values include:* **human** the content was 'naturally' typed or spoken by a human user

**template** the content was generated after a human user activated a template for its insertion

**system** the content was generated by the system, i.e. the CMC environment

**bot** the content was generated by a bot, i.e. a non-human agent, typically one that is not part of the CMC environment itself

**unspecified** the content was generated by an unknown or unspecified process

automatic system message in chat: user moves on to another chatroom

```
<post  type="event"
  generatedBy="system"
  who="#system"
  rend="color:blue">
  <p>
    <name  type="nickname"
      corresp="#A02">McMike</name> geht
     in einen anderen Raum:
<name  type="roomname">Kreuzfahrt</name>
  </p>
</post>
```

automatic system message in chat: user enters a chatroom

```
<post  type="event"
  generatedBy="system">
  <p>
    <name  type="nickname"
      corresp="#A08">c_bo</name> betritt
     den Raum. </p>
</post>
```

automatic system message in chat: user changes his font color

```
<post  type="event"
  generatedBy="system"
  rend="color:red">
  <p>
    <name  type="nickname"
      corresp="#A08">c_bo</name> hat die
     Farbe gewechselt.
  </p>
</post>
```

An automatic signature of user including an automatic timestamp (Wikipedia discussion, anonymized). The specification of *generatedBy* at the inner element <signed> is meant to override the specification at the outer element <post>. This is generally possible when the outer *generatedBy* value is "human".

```
<post  type="standard"
  generatedBy="human"
  indentLevel="2"
  synch="#t00394407"
  who="#WU00005582">
  <p> Kurze Nachfrage: Die Hieros für den Goldnamen stammen
     auch von Beckerath gem. Literatur ? Grüße --</p>
  <signed  generatedBy="template"
    rend="inline">
    <gap  reason="signatureContent"/>
    <time  generatedBy="template">18:50, 22. Okt. 2008
(CEST)</time>
  </signed>
</post>
```

Wikipedia talk page: user signature

```
<post  type="written"
  generatedBy="human">
```

```
<!-- ... main content of posting ... -->
  <signed generatedBy="template">
    <gap reason="signatureContent"/>
    <time generatedBy="template">12:01, 12. Jun. 2009
(CEST)</time>
  </signed>
</post>
```

---

**att.combinable** provides attributes indicating how multiple references to the same object in a schema should be combined.

*Module* tagdocs — Documentation Elements

*Members* att.identified[attDef classSpec constraintSpec dataSpec elementSpec macroSpec moduleSpec paramSpec schemaSpec] defaultVal remarks valDesc valItem valList

*Attributes*
- att.deprecated
  - *@validUntil*

`@mode` specifies the effect of this declaration on its parent object.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **add** this declaration is part of the current definitions.[Default]

**delete** the object with the same identifier (and namespace in the case of an element or attribute) as the current one is no longer part of the current definitions

**change** this declaration changes the object of the same identifier (and namespace in the case of an element or attribute) in the current definition

**replace** this declaration replaces the object with the same identifier (and namespace in the case of an element or attribute) in the current definition.

An ODD processor should handle the values for *mode* as follows:

*Note* **add** the object should be created (processing any children in add mode); raise an error if the object has an identifier and an object with the same identifier (and namespace in the case of an element or attribute) already exists.

**replace** use this object in preference to any existing object with the same identifier (and namespace in the case of an element or attribute), and ignore any children of that object; process any new children in replace mode

**delete** do not process this object or any existing object with the same identifier (and namespace in the case of an element or attribute); the element carrying this attribute must be empty: raise an error if there are any children

**change** process this object and its children, replacing only those components of an existing object of the same type with the same identifier (and namespace in case of an element or attribute) that are explicitly specified. This will change only those existing attribute values and child elements that appear both in the original specification and in the new specification. Other children from the original specification will be preserved.

---

**att.coordinated** provides attributes that can be used to position their parent element within a two dimensional coordinate system.

*Module* transcr — Representation of Primary Sources

*Members* line path surface zone

*Attributes*

`@start` indicates the element within a transcription of the text containing at least the start of the writing represented by this zone or surface.

*Status* Optional

*Datatype* teidata.pointer

`@ulx` gives the x coordinate value for the upper left corner of a rectangular space.
*Status* Optional

*Datatype* teidata.numeric

`@uly` gives the y coordinate value for the upper left corner of a rectangular space.
*Status* Optional

*Datatype* teidata.numeric

`@lrx` gives the x coordinate value for the lower right corner of a rectangular space.
*Status* Optional

*Datatype* teidata.numeric

`@lry` gives the y coordinate value for the lower right corner of a rectangular space.
*Status* Optional

*Datatype* teidata.numeric

`@points` identifies a two dimensional area by means of a series of pairs of numbers, each of which gives the x,y coordinates of a point on a line enclosing the area.
*Status* Optional

*Datatype* 3–∞ occurrences of teidata.point separated by whitespace

---

**att.cReferencing** provides attributes that may be used to supply a *canonical reference* as a means of identifying the target of a pointer.

*Module* tei — The TEI Infrastructure

*Members* gloss ptr ref term

*Attributes*

`@cRef` (canonical reference) specifies the destination of the pointer by supplying a canonical reference expressed using the scheme defined in a <refsDecl> element in the TEI header.
*Status* Optional

*Datatype* teidata.text

*Note* The value of *cRef* should be constructed so that when the algorithm for the resolution of canonical references (described in section *17.2.5. Canonical References*) is applied to it the result is a valid URI reference to the intended target.

The <refsDecl> to use may be indicated with the *decls* attribute.

Currently these Guidelines only provide for a single canonical reference to be encoded on any given <ptr> element.

---

**att.damaged** provides attributes describing the nature of any physical damage affecting a reading. [*12.3.3.1. Damage, Illegibility, and Supplied Text 1.3.1. Attribute Classes*]

*Module* tei — The TEI Infrastructure

*Members* damage damageSpan

*Attributes*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    - *@atLeast*
    - *@atMost*
    - *@min*
    - *@max*
    - *@confidence*
- att.written
  - *@hand*

`@agent` categorizes the cause of the damage, if it can be identified.
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **rubbing** damage results from rubbing of the leaf edges
**mildew** damage results from mildew on the leaf surface
**smoke** damage results from smoke

`@degree` provides a coded representation of the degree of damage, either as a number between 0 (undamaged) and 1 (very extensively damaged), or as one of the codes high, medium, low, or unknown. The <damage> element with the *degree* attribute should only be used where the text may be read with some confidence; text supplied from other sources should be tagged as <supplied>.
*Status* Optional
*Datatype* teidata.probCert
*Note* The <damage> element is appropriate where it is desired to record the fact of damage although this has not affected the readability of the text, for example a weathered inscription. Where the damage has rendered the text more or less illegible either the <unclear> tag (for partial illegibility) or the <gap> tag (for complete illegibility, with no text supplied) should be used, with the information concerning the damage given in the attribute values of these tags. See section *12.3.3.2. Use of the gap, del, damage, unclear, and supplied Elements in Combination* for discussion of the use of these tags in particular circumstances.

`@group` assigns an arbitrary number to each stretch of damage regarded as forming part of the same physical phenomenon.
*Status* Optional
*Datatype* teidata.count

---

## att.datable provides attributes for normalization of elements that contain dates, times, or datable events. [*3.6.4. Dates and Times 14.4. Dates*]

*Module* tei — The TEI Infrastructure

*Members* att.gaijiProp[localProp unicodeProp unihanProp] acquisition affiliation age altIdentifier application author binding birth bloc change climate conversion country creation custEvent date death district docDate editor education event eventName faith floruit funder gender geogFeat geogName idno langKnowledge langKnown licence location mapping meeting name nationality objectName occupation offset orgName origDate origPlace origin persName persPronouns placeName population post precision principal provenance region relation residence resp seal settlement sex socecStatus sponsor stamp state terrain time title trait unitDecl unitDef

*Attributes*
- att.datable.custom
  - *@when-custom*

- *@notBefore-custom*
- *@notAfter-custom*
- *@from-custom*
- *@to-custom*
- *@datingPoint*
- *@datingMethod*

• att.datable.iso
- *@when-iso*
- *@notBefore-iso*
- *@notAfter-iso*
- *@from-iso*
- *@to-iso*

• att.datable.w3c
- *@when*
- *@notBefore*
- *@notAfter*
- *@from*
- *@to*

`@period` supplies pointers to one or more definitions of named periods of time (typically <category>s, <date>s, or <event>s) within which the datable item is understood to have occurred.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

This 'superclass' provides attributes that can be used to provide normalized values of temporal information. By default, the attributes from the `att.datable.w3c` class are provided. If the module for names & dates is loaded, this class also provides attributes from the `att.datable.iso` and `att.datable.custom` classes. In general, the possible values of attributes restricted to the W3C datatypes form a subset of those values available via the ISO 8601 standard. However, the greater expressiveness of the ISO datatypes may not be needed, and there exists much greater software support for the W3C datatypes.

---

## att.datable.custom

Module provides attributes for normalization of elements that contain datable events to a custom dating system (i.e. other than the Gregorian used by W3 and ISO). [*14.4. Dates*]

*Module* namesdates — Names, Dates, People, and Places

*Members* att.datable[att.gaijiProp[localProp unicodeProp unihanProp] acquisition affiliation age altIdentifier application author binding birth bloc change climate conversion country creation custEvent date death district docDate editor education event eventName faith floruit funder gender geogFeat geogName idno langKnowledge langKnown licence location mapping meeting name nationality objectName occupation offset orgName origDate origPlace origin persName persPronouns placeName population post precision principal provenance region relation residence resp seal settlement sex socecStatus sponsor stamp state terrain time title trait unitDecl unitDef]

*Attributes*

`@when-custom` supplies the value of a date or time in some custom standard form.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

The following are examples of custom date or time formats that are *not* valid ISO or W3C format normalizations, normalized to a different dating system

```
<p>Alhazen died in Cairo on the
<date when="1040-03-06"
   when-custom="431-06-12"> 12th day of Jumada t-Tania, 430 AH
  </date>.</p>
<p>The current world will end at the
<date when="2012-12-21"
   when-custom="13.0.0.0.0">end of B'ak'tun 13</date>.</p>
<p>The Battle of Meggidu
(<date when-custom="Thutmose_III:23">23rd year of reign of
Thutmose III</date>).</p>
<p>Esidorus bixit in pace annos LXX plus minus sub
<date when-custom="Ind:4-10-11">die XI mensis Octobris
indictione IIII</date>
</p>Not all custom date formulations will have Gregorian
equivalents.The when-custom attribute and other custom dating
are not constrained to a datatype by the TEI, but individual
projects are recommended to regularize and document their dating
formats.
```

`@notBefore-custom` specifies the earliest possible date for the event in some custom standard form.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

`@notAfter-custom` specifies the latest possible date for the event in some custom standard form.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

`@from-custom` indicates the starting point of the period in some custom standard form.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

```
<event xml:id="FIRE1"
  datingMethod="#julian"
  from-custom="1666-09-02"
  to-custom="1666-09-05">
  <head>The Great Fire of London</head>
  <p>The Great Fire of London burned through a large part
     of the city of London.</p>
</event>
```

`@to-custom` indicates the ending point of the period in some custom standard form.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

`@datingPoint` supplies a pointer to some location defining a named point in time with reference to which the datable item is understood to have occurred.

*Status* Optional

*Datatype* teidata.pointer

`@datingMethod` supplies a pointer to a <calendar> element or other means of interpreting the values of the custom dating attributes.

*Status* Optional

*Datatype* teidata.pointer

```
Contayning the Originall, Antiquity, Increaſe, Moderne
eſtate, and deſcription of that Citie, written in the yeare
<date when-custom="1598"
  calendar="#julian"
  datingMethod="#julian">1598</date>. by Iohn Stow
```

Citizen of London.In this example, the *calendar* attribute points to a <calendar> element for the Julian calendar, specifying that the text content of the <date> element is a Julian date, and the *datingMethod* attribute also points to the Julian calendar to indicate that the content of the *when-custom* attribute value is Julian too.

```
<date when="1382-06-28"
  when-custom="6890-06-20"
  datingMethod="#creationOfWorld"> μηνὶ Ἰουνίου εἰς <num>κ</num>
ἔτους <num>ϛωϡ</num>
</date>In this example, a date is given in a Mediaeval text
measured 'from the creation of the world', which is normalized
(in when) to the Gregorian date, but is also normalized (in
when-custom) to a machine-actionable, numeric version of the
date from the Creation.
```

*Note* Note that the *datingMethod* attribute (unlike *calendar* defined in att.datable) defines the calendar or dating system to which the date described by the parent element is normalized (i.e. in the *when-custom* or other *X-custom* attributes), *not* the calendar of the original date in the element.

---

## att.datable.iso provides attributes for normalization of elements that contain datable events using the ISO 8601:2004 standard. [*3.6.4. Dates and Times 14.4. Dates*]

*Module* namesdates — Names, Dates, People, and Places

*Members* att.datable[att.gaijiProp[localProp unicodeProp unihanProp] acquisition affiliation age altIdentifier application author binding birth bloc change climate conversion country creation custEvent date death district docDate editor education event eventName faith floruit funder gender geogFeat geogName idno langKnowledge langKnown licence location mapping meeting name nationality objectName occupation offset orgName origDate origPlace origin persName persPronouns placeName population post precision principal provenance region relation residence resp seal settlement sex socecStatus sponsor stamp state terrain time title trait unitDecl unitDef]

*Attributes*

`@when-iso` supplies the value of a date or time in a standard form.

*Status* Optional

*Datatype* teidata.temporal.iso

The following are examples of ISO date, time, and date & time formats that are *not* valid W3C format normalizations.

```
<date when-iso="1996-09-24T07:25+00">Sept. 24th, 1996 at 3:25
in the morning</date>
<date when-iso="1996-09-24T03:25-04">Sept. 24th, 1996 at 3:25
in the morning</date>
<time when-iso="1999-01-04T20:42-05">4 Jan 1999 at 8:42
pm</time>
<time when-iso="1999-W01-1T20,70-05">4 Jan 1999 at 8:42
pm</time>
<date when-iso="2006-05-18T10:03">a few minutes after ten in
the morning on Thu 18 May</date>
<time when-iso="03:00">3 A.M.</time>
<time when-iso="14">around two</time>
<time when-iso="15,5">half past three</time>All of the examples
of the when attribute in the att.datable.w3c class are also
valid with respect to this attribute.
```

```
He likes to be punctual. I said <q>
  <time when-iso="12">around noon</time>
```

```
</q>, and he showed up at <time  when-iso="12:00:00">12
O'clock</time> on the dot.The second occurrence of <time> could
have been encoded with the when attribute, as 12:00:00 is a
valid time with respect to the W3C XML Schema Part 2: Datatypes
Second Edition specification. The first occurrence could not.
```

`@notBefore-iso` specifies the earliest possible date for the event in standard form, e.g. yyyy-mm-dd.
*Status* Optional
*Datatype* teidata.temporal.iso

`@notAfter-iso` specifies the latest possible date for the event in standard form, e.g. yyyy-mm-dd.
*Status* Optional
*Datatype* teidata.temporal.iso

`@from-iso` indicates the starting point of the period in standard form.
*Status* Optional
*Datatype* teidata.temporal.iso

`@to-iso` indicates the ending point of the period in standard form.
*Status* Optional
*Datatype* teidata.temporal.iso

The value of these attributes should be a normalized representation of the date, time, or combined date & time intended, in any of the standard formats specified by ISO 8601:2004, using the Gregorian calendar.

If both *when-iso* and *dur-iso* are specified, the values should be interpreted as indicating a span of time by its starting time (or date) and duration. That is,

```
<date  when-iso="2007-06-01"  dur-iso="P8D"/>
```

indicates the same time period as

```
<date  when-iso="2007-06-01/P8D"/>
```

In providing a 'regularized' form, no claim is made that the form in the source text is incorrect; the regularized form is simply that chosen as the main form for purposes of unifying variant forms under a single heading.

---

## att.datable.w3c

`att.datable.w3c` provides attributes for normalization of elements that contain datable events conforming to the W3C XML Schema Part 2: Datatypes Second Edition. [*3.6.4. Dates and Times 14.4. Dates*]

*Module* tei — The TEI Infrastructure

*Members* att.datable[att.gaijiProp[localProp unicodeProp unihanProp] acquisition affiliation age altIdentifier application author binding birth bloc change climate conversion country creation custEvent date death district docDate editor education event eventName faith floruit funder gender geogFeat geogName idno langKnowledge langKnown licence location mapping meeting name nationality objectName occupation offset orgName origDate origPlace origin persName persPronouns placeName population post precision principal provenance region relation residence resp seal settlement sex socecStatus sponsor stamp state terrain time title trait unitDecl unitDef]

*Attributes*

`@when` supplies the value of the date or time in a standard form, e.g. yyyy-mm-dd.
*Status* Optional
*Datatype* teidata.temporal.w3c

Examples of W3C date, time, and date & time formats. `<p>`
`<date  when="1945-10-24">24 Oct 45</date>`

```
  <date  when="1996-09-24T07:25:00Z">September 24th, 1996 at 3:25
in the morning</date>
  <time  when="1999-01-04T20:42:00-05:00">Jan 4 1999 at 8
pm</time>
  <time  when="14:12:38">fourteen twelve and 38 seconds</time>
  <date  when="1962-10">October of 1962</date>
  <date  when="--06-12">June 12th</date>
  <date  when="---01">the first of the month</date>
  <date  when="--08">August</date>
  <date  when="2006">MMVI</date>
  <date  when="0056">AD 56</date>
  <date  when="-0056">56 BC</date>
</p>
This list begins in
the year 1632, more precisely on Trinity Sunday, i.e. the Sunday
after
Pentecost, in that year the
<date  calendar="#julian"
  when="1632-06-06">27th of May (old style)</date>.
<opener>
  <dateline>
    <placeName>Dorchester, Village,</placeName>
    <date  when="1828-03-02">March 2d. 1828.</date>
  </dateline>
  <salute>To
     Mrs. Cornell,</salute> Sunday
<time  when="12:00:00">noon.</time>
</opener>
```

`@notBefore` specifies the earliest possible date for the event in standard form, e.g. yyyy-mm-dd.
*Status* Optional
*Datatype* teidata.temporal.w3c

`@notAfter` specifies the latest possible date for the event in standard form, e.g. yyyy-mm-dd.
*Status* Optional
*Datatype* teidata.temporal.w3c

`@from` indicates the starting point of the period in standard form, e.g. yyyy-mm-dd.
*Status* Optional
*Datatype* teidata.temporal.w3c

`@to` indicates the ending point of the period in standard form, e.g. yyyy-mm-dd.
*Status* Optional
*Datatype* teidata.temporal.w3c

*Schematron* <sch:rule context="tei:*[@when]"> <sch:report test="@notBefore|@notAfter|@from|@to" role="nonfatal">The @when attribute cannot be used with any other att.datable.w3c attributes.</sch:report> </sch:rule>

*Schematron* <sch:rule context="tei:*[@from]"> <sch:report test="@notBefore" role="nonfatal">The @from and @notBefore attributes cannot be used together.</sch:report> </sch:rule>

*Schematron* <sch:rule context="tei:*[@to]"> <sch:report test="@notAfter" role="nonfatal">The @to and @notAfter attributes cannot be used together.</sch:report> </sch:rule>

*Example*

```
<date  from="1863-05-28"  to="1863-06-01">28 May through 1 June 1863</date>
```

*Note* The value of these attributes should be a normalized representation of the date, time, or combined date & time intended, in any of the standard formats specified by XML Schema Part 2: Datatypes Second Edition, using the Gregorian calendar.

The most commonly-encountered format for the date portion of a temporal attribute is `yyyy-mm-dd`, but `yyyy`, `--mm`, `---dd`, `yyyy-mm`, or `--mm-dd` may also be used. For the time part, the form `hh:mm:ss` is used.

Note that this format does not currently permit use of the value 0000 to represent the year 1 BCE; instead the value -0001 should be used.

---

**att.datcat** provides attributes that are used to align XML elements or attributes with the appropriate Data Categories (DCs) defined by an external taxonomy, in this way establishing the identity of information containers and values, and providing means of interpreting them. [*10.5.2. Lexical View 19.3. Other Atomic Feature Values*]

*Module* tei — The TEI Infrastructure

*Members* att.lexicographic[case colloc def entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number oRef orth pRef per pos pron re sense subc syll tns usg xr] att.segLike[c cl m pc phr s seg w] binary category f fDecl fs fsDecl numeric string symbol tagUsage taxonomy

*Attributes*

`@datcat` provides a pointer to a definition of, and/or general information about, (a) an information container (element or attribute) or (b) a value of an information container (element content or attribute value), by referencing an external taxonomy or ontology. If *valueDatcat* is present in the immediate context, this attribute takes on role (a), while *valueDatcat* performs role (b).
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

`@valueDatcat` provides a definition of, and/or general information about a value of an information container (element content or attribute value), by reference to an external taxonomy or ontology. Used especially where a contrast with *datcat* is needed.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

`@targetDatcat` provides a definition of, and/or general information about, information structure of an object referenced or modeled by the containing element, by reference to an external taxonomy or ontology. This attribute has the characteristics of the *datcat* attribute, except that it addresses not its containing element, but an object that is being referenced or modeled by its containing element.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Example* The example below presents the TEI encoding of the name-value pair `<part of speech, common noun>`, where the name (key) 'part of speech' is abbreviated as 'POS', and the value, 'common noun' is symbolized by 'NN'. The entire name-value pair is encoded by means of the element <f>. In TEI XML, that element acts as the container, labeled with the *name* attribute. Its contents may be complex or simple. In the case at hand, the content is the symbol 'NN'.The *datcat* attribute relates the feature *name* (i.e., the key) to the data category 'part of speech', while the attribute *valueDatcat* relates the feature *value* to the data category *common noun*. Both these data categories should be defined in an external and preferably open reference taxonomy or ontology.

```
<fs>
  <f  name="POS"
    datcat="http://hdl.handle.net/11459/CCR_C-396_5a972b93-2294-ab5c-a541-7c344c5f26c3">
    <symbol  valueDatcat="http://hdl.handle.net/11459/CCR_C-1256_7ec6083c-23d4-224d-6f94-eecbe
      value="NN"/>
  </f>
<!-- ... -->
</fs>
```

'NN' is the symbol for common noun used e.g. in the CLAWS-7 tagset defined by the University Centre

for Computer Corpus Research on Language at the University of Lancaster. The very same data category used for tagging an early version of the British National Corpus, and coming from the BNC Basic (C5) tagset, uses the symbol 'NN0' (rather than 'NN'). Making these values semantically interoperable would be extremely difficult without a human expert if they were not anchored in a single point of an established reference taxonomy of morphosyntactic data categories. In the case at hand, the string 'http://hdl.handle.net/11459/CCR_C-1256_7ec6083c-23d4-224d-6f94-eecbe6861545' is both a persistent identifier of the data category in question, as well as a pointer to a shared definition of *common noun*.While the symbols 'NN', 'NN0', and many others (often coming from languages other than English) are implicitly members of the container category 'part of speech', it is sometimes useful not to rely on such an implicit relationship but rather use an explicit identifier for that data category, to distinguish it from other morphosyntactic data categories, such as gender, tense, etc. For that purpose, the above example uses the *datcat* attribute to reference a definition of *part of speech*. The reference taxonomy in this example is the CLARIN Concept Registry.If the feature structure markup exemplified above is to be repeated many times in a single document, it is much more efficient to gather the persistent identifiers in a single place and to only reference them, implicitly or directly, from feature structure markup. The following example is much more concise than the one above and relies on the concepts of feature structure declaration and feature value library, discussed in chapter *19. Feature Structures*.

```
<fs>
  <f  name="POS"  fVal="#commonNoun"/>
<!-- ... -->
</fs>
```

The assumption here is that the relevant feature values are collected in a place that the annotation document in question has access to — preferably, a single document per linguistic resource, for example an <fsdDecl> that is XIncluded as a sibling of <text> or a child of <encodingDesc>; a <taxonomy> available resource-wide (e.g., in a shared header) is also an option.The example below presents an <fvLib> element that collects the relevant feature values (most of them omitted). At the same time, this example shows one way of encoding a *tagset*, i.e., an established inventory of values of (in the case at hand) morphosyntactic categories.

```
<fvLib  n="POS values">
  <symbol  xml:id="commonNoun"  value="NN"
    datcat="http://hdl.handle.net/11459/CCR_C-396_5a972b93-2294-ab5c-a541-7c344c5f26c3"/>
  <symbol  xml:id="properNoun"  value="NP"
    datcat="http://hdl.handle.net/11459/CCR_C-1371_fbebd9ec-a7f4-9a36-d6e9-88ee16b944ae"/>
<!-- ... -->
</fvLib>
```

Note that these Guidelines do not prescribe a specific choice between *datcat* and *valueDatcat* in such cases. The former is the generic way of referencing a data category, whereas the latter is more specific, in that it references a data category that represents a value. The choice between them comes into play where a single element — or a tight element complex, such as the <f>/<symbol> complex illustrated above — make it necessary or useful to distinguish between the container data category and its value.

*Example* In the context of dictionaries designed with semantic interoperability in mind, the following example ensures that the <pos> element is interpreted as the same information container as in the case of the example of <f name="POS"> above.

```
<gramGrp>
  <pos  datcat="http://hdl.handle.net/11459/CCR_C-396_5a972b93-2294-ab5c-a541-7c344c5f26c3"
    valueDatcat="http://hdl.handle.net/11459/CCR_C-1256_7ec6083c-23d4-224d-6f94-eecbe6861545">
</gramGrp>
```

Efficiency of this type of interoperable markup demands that the references to the particular data categories should best be provided in a single place within the dictionary (or a single place within the project), rather than being repeated inside every entry. For the container elements, this can be achieved at the level of <tagUsage>, although here, the *valueDatcat* attribute should be used, because it is not the

<tagUsage> element that is associated with the relevant data category, but rather the element <pos> (or <case>, etc.) that is described by <tagUsage>:

```
<tagsDecl  partial="true">
<!-- ... -->
  <namespace  name="http://www.tei-c.org/ns/1.0">
    <tagUsage  gi="pos"
      targetDatcat="http://hdl.handle.net/11459/CCR_C-396_5a972b93-2294-ab5c-a541-7c344c5f26c3
the part of speech.</tagUsage>
    <tagUsage  gi="case"
      targetDatcat="http://hdl.handle.net/11459/CCR_C-1840_9f4e319c-f233-6c90-9117-7270e215f03
information about the grammatical case that the described form is inflected
for.</tagUsage>
<!-- ... -->
  </namespace>
</tagsDecl>
```

Another possibility is to shorten the URIs by means of the <prefixDef> mechanism, as illustrated below:

```
<listPrefixDef>
  <prefixDef  ident="ccr"  matchPattern="pos"
    replacementPattern="http://hdl.handle.net/11459/CCR_C-396_5a972b93-2294-ab5c-a541-7c344c5f
  <prefixDef  ident="ccr"  matchPattern="adj"
    replacementPattern="http://hdl.handle.net/11459/CCR_C-1230_23653c21-fca1-edf8-fd7c-3df2d64
</listPrefixDef>
<!-- ... -->
<entry>
<!--...-->
  <form>
    <orth>isotope</orth>
  </form>
  <gramGrp>
    <pos  datcat="ccr:pos"
      valueDatcat="ccr:adj">adj</pos>
  </gramGrp>
<!--...-->
</entry>
```

This mechanism creates implications that are not always wanted, among others, in the case at hand, suggesting that the identifiers ‘pos’ and ‘adj’ belong to a namespace associated with the CLARIN Concept Repository (CCR), whereas that is solely a shorthand mechanism whose scope is the current resource. Documenting this clearly in the header of the dictionary is therefore advised.Yet another possibility is to associate the information about the relationship between a TEI markup element and the data category that it is intended to model already at the level of modeling the dictionary resource, that is, at the level of the ODD, in the <equiv> element that is a child of <elementSpec> or <attDef>.

*Example* The <taxonomy> element is a handy tool for encoding taxonomies that are later referenced by `att.datcat` attributes, but it can also act as an intermediary device, for example holding a fragment of an external taxonomy (or ‘flattening’ an external ontology) that is relevant to the project or document at hand. (It is also imaginable that, for the purpose of the project at hand, the local <taxonomy> element combines vocabularies that originate from more than one external taxonomy or ontology.) In such cases, the <taxonomy> creates a local layer of indirection: the `att.datcat` attributes internal to the resource may reference the <category> elements stored in the header (as well as the <taxonomy> element itself), whereas these same <category> and <taxonomy> elements use `att.datcat` attributes to reference the original taxonomy or ontology.

```
<encodingDesc>
<!-- ... -->
  <classDecl>
<!-- ... -->
    <taxonomy  xml:id="UD-SYN"
      datcat="https://universaldependencies.org/u/dep/index.html">
      <desc>
        <term>UD syntactic relations</term>
      </desc>
      <category  xml:id="acl"
        valueDatcat="https://universaldependencies.org/u/dep/acl.html">
        <catDesc>
          <term>acl</term>: Clausal modifier of noun (adjectival
clause)</catDesc>
      </category>
      <category  xml:id="acl_relcl"
        valueDatcat="https://universaldependencies.org/u/dep/acl-relcl.html">
        <catDesc>
          <term>acl:relcl</term>: relative clause modifier</catDesc>
      </category>
      <category  xml:id="advcl"
        valueDatcat="https://universaldependencies.org/u/dep/advcl.html">
        <catDesc>
          <term>advcl</term>: Adverbial clause modifier</catDesc>
      </category>
<!-- ... -->
    </taxonomy>
  </classDecl>
</encodingDesc>
```

The above fragment was excerpted from the GB subset of the ParlaMint project in April 2023, and enriched with `att.datcat` attributes for the purpose of illustrating the mechanism described here.Note that, in the ideal case, the values of `att.datcat` attributes should be persistent identifiers, and that the addressing scheme of Universal Dependencies is treated here as persistent for the sake of illustration. Note also that the contrast between *datcat* used on <taxonomy> on the one hand, and the *valueDatcat* used on <category> on the other, is not mandatory: both kinds of relations could be encoded by means of the generic *datcat* attribute, but using the former for the container and the latter for the content is more user-friendly.

*Example* The *targetDatcat* attribute is designed to be used in, e.g., feature structure declarations, and is analogous to the *targetLang* attribute of the `att.pointing` class, in that it describes the object that is being referenced, rather than the referencing object.

```
<fDecl  name="POS"
  targetDatcat="http://hdl.handle.net/11459/CCR_C-396_5a972b93-2294-ab5c-a541-7c344c5f26c3">
  <fDescr>part of speech (morphosyntactic category)</fDescr>
  <vRange>
    <vAlt>
      <symbol  value="NN"
        datcat="http://hdl.handle.net/11459/CCR_C-1256_7ec6083c-23d4-224d-6f94-eecbe6861545"/>
      <symbol  value="NP"
        datcat="http://hdl.handle.net/11459/CCR_C-1371_fbebd9ec-a7f4-9a36-d6e9-88ee16b944ae"/>
<!-- ... -->
    </vAlt>
  </vRange>
</fDecl>
```

Above, the <fDecl> uses *targetDatcat*, because if it were to use *datcat*, it would be asserting that it is an

instance of the container data category *part of speech*, whereas it is not — it models a container (<f>) that encodes a part of speech. Note also that it is the <f> that is modeled above, not its values, which are used as direct references to data categories; hence the use of *datcat* in the <symbol> element.

*Example* The `att.datcat` attributes can be used for any sort of taxonomies. The example below illustrates their usefulness for describing usage domain labels in dictionaries on the example of the *Diccionario da Lingua Portugueza* by António de Morais Silva, retro-digitised in the MORDigital project.

```
<!-- in the dictionary header --><encodingDesc>
  <classDecl>
    <taxonomy  xml:id="domains">
<!--...-->
     <category  xml:id="domain.medical_and_health_sciences">
       <catDesc  xml:lang="en">Medical and Health Sciences</catDesc>
       <catDesc  xml:lang="pt">Ciências Médicas e da Saúde</catDesc>
       <category  xml:id="domain.medical_and_health_sciences.medicine"
         valueDatcat="https://vocabs.rossio.fcsh.unl.pt/pub/morais_domains/pt/page/0025">
         <catDesc  xml:lang="en">
           <term>Medicine</term>
           <gloss>
<!--...-->
           </gloss>
         </catDesc>
         <catDesc  xml:lang="pt">
           <term>Medicina</term>
           <gloss>
<!--...-->
           </gloss>
         </catDesc>
       </category>
     </category>
<!--...-->
    </taxonomy>
  </classDecl>
</encodingDesc>
<!--
inside an <entry> element: -->
<usg  type="domain"
  valueDatcat="#domain.medical_and_health_sciences.medicine">Med.</usg>
```

In the Morais dictionary, the relevant domain labels are in the header, getting referenced inside the dictionary, from <usg> elements. The vocabulary used for dictionary-internal labelling is in turn anchored in the MorDigital controlled vocabulary service of the NOVA University of Lisbon – School of Social Sciences and Humanities (NOVA FCSH).

*Note* The TEI Abstract Model can be expressed as a hierarchy of attribute-value matrices (AVMs) of various types and of various levels of complexity, nested or grouped in various ways. At the most abstract level, an AVM consists of an information container and the value (contents) of that container.

A simple example of an XML serialization of such structures is, on the one hand, the opening and closing tags that delimit and name the container, and, on the other, the content enclosed by the two tags that constitues the value. An analogous example is an attribute name and the value of that attribute.

In a TEI XML example of two equivalent serializations expressing the name-value pair `<part-of-speech,common-noun>`, namely `<pos>commonNoun</pos>` and `pos="common-noun"`, one would classify the element <pos> and the attribute *pos* as containers (mapping onto the first member of the relevant name-value pair), while the character data content of <pos> or the value of *pos* would be seen as mapping onto the second member of the pair.

The `att.datcat` class provides means of addressing the containers and their values, while at the same time providing a way to interpret them in the context of external taxonomies or ontologies.

Aligning e.g. both the <pos> element and the *pos* attribute with the same value of an external reference point (i.e., an entry in an agreed taxonomy) affirms the identity of the concept serialised by both the element container and the attribute container, and optionally provides a definition of that concept (in the case at hand, the concept *part of speech*).

The value of the `att.datcat` attributes should be a PID (persistent identifier) that points to a specific — and, ideally, shared — taxonomy or ontology. Among the resources that can, to a lesser or greater extent, be used as inventories of (more or less) standardized linguistic categories are the GOLD ontology, CLARIN CCR, OLiA, or TermWeb's DatCatInfo, and also the Universal Dependencies inventory, on the assumption that its URIs are going to persist. It is imaginable that a project may choose to address a local taxonomy store instead, but this risks losing the advantage of interchangeability with other projects.

Historically, *datcat* and *valueDatcat* originate from the (now obsolete) ISO 12620:2009 standard, describing the data model and procedures for a Data Category Registry (DCR). The current version of that standard, ISO 12620-1, does not standardize the serialization of pointers, merely mentioning the TEI `att.datcat` as an example.

Note that no constraint prevents the occurrence of a combination of `att.datcat` attributes: the <fDecl> element, which is a natural bearer of the *targetDatcat* attribute, is an instance of a specific modeling element, and, in principle, could be semantically fixed by an appropriate reference taxonomy of modeling devices.

---

**att.declarable** provides attributes for those elements in the TEI header which may be independently selected by means of the special purpose *decls* attribute. [*16.3. Associating Contextual Information with a Text*]

*Module* tei — The TEI Infrastructure

*Members* availability bibl biblFull biblStruct broadcast correction correspDesc editorialDecl equipment geoDecl hyphenation interpretation langUsage listApp listBibl listEvent listNym listObject listOrg listPerson listPlace metDecl normalization particDesc projectDesc punctuation quotation recording refsDecl samplingDecl scriptStmt segmentation seriesStmt settingDesc sourceDesc stdVals styleDefDecl textClass textDesc xenoData

*Attributes*

`@default` indicates whether or not this element is selected by default when its parent is selected.

*Status* Optional

*Datatype* teidata.truthValue

*Legal values are:* **true** This element is selected if its parent is selected

**false** This element can only be selected explicitly, unless it is the only one of its kind, in which case it is selected if its parent is selected.[Default]

*Schematron* <sch:pattern id="declarable" abstract="true"> <!– parameter 'tde' is for "this declarable element (type)" –> <sch:rule context="$tde[ ancestor::tei:teiHeader and following-sibling::$tde and not( preceding-sibling::$tde ) ]"> <sch:report test="../child::$tde[ not( @xml:id ) ]"> When there is more than one <sch:name/>, each must have an @xml:id </sch:report> <sch:assert test="count( ../child::$tde[ normalize-space( @default ) = ('1','true') ] ) eq 1"> When there is more than one <sch:name/>, one and only one must have a @default of 'true'. </sch:assert> </sch:rule> </sch:pattern>

*Note* The rules governing the association of declarable elements with individual parts of a TEI text are fully defined in chapter *16.3. Associating Contextual Information with a Text.* Only one element of a particular type may have a *default* attribute with a value of true.

---

**att.declaring** provides attributes for elements which may be independently associated with a particular declarable element within the header, thus overriding the inherited default for that element. [*16.3. Associating Contextual Information with a Text*]

*Module* tei — The TEI Infrastructure

*Members* ab back body div div1 div2 div3 div4 div5 div6 div7 facsimile floatingText front geo gloss graphic group lg listAnnotation media msDesc object p ptr ref sourceDoc standOff surface surfaceGrp term text u

*Attributes*

`@decls` (declarations) identifies one or more *declarable elements* within the header, which are understood to apply to the element bearing this attribute and its content.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

The rules governing the association of declarable elements with individual parts of a TEI text are fully defined in chapter *16.3. Associating Contextual Information with a Text*.

---

## att.deprecated provides attributes indicating how a deprecated feature will be treated in future releases.

*Module* tagdocs — Documentation Elements

*Members* att.combinable[att.identified[attDef classSpec constraintSpec dataSpec elementSpec macroSpec moduleSpec paramSpec schemaSpec] defaultVal remarks valDesc valItem valList]

*Attributes*

`@validUntil` provides a date before which the construct being defined will not be removed.

*Status* Optional

*Datatype* date

*Schematron* <sch:rule context="tei:*[@validUntil]"> <sch:let name="advance_warning_period" value="current-date() + xs:dayTimeDuration('P60D')"/> <sch:let name="me_phrase" value="if (@ident) then concat('The ', @ident ) else concat('This ', local-name(.), ' of ', ancestor::tei:*[@ident][1]/@ident )"/> <sch:assert test="@validUntil cast as xs:date ge current-date()"> <sch:value-of select="concat( $me_phrase, ' construct is outdated (as of ', @validUntil, '); ODD processors may ignore it, and its use is no longer supported' )"/> </sch:assert> <sch:assert role="warning" test="@validUntil cast as xs:date ge $advance_warning_period"> <sch:value-of select="concat( $me_phrase, ' construct becomes outdated on ', @validUntil )"/> </sch:assert> </sch:rule>

*Schematron* <sch:rule context="tei:*[@validUntil][ not( self::tei:valDesc | self::tei:valList | self::tei:defaultVal | self::tei:remarks )]"> <sch:assert test="child::tei:desc[ @type eq 'deprecationInfo']"> A deprecated construct should include, whenever possible, an explanation, but this <sch:value-of select="name(.)"/> does not have a child <desc type="deprecationInfo"></sch:assert> </sch:rule>

*Note* The value of this attribute should represent a date (in standard `yyyy-mm-dd` format) which is later than the date on which the attribute is added to an ODD. Technically, this attribute asserts only the intent to leave a construct in future releases of the markup language being defined up to at least the specified date, and makes no assertion about what happens past that date. In practice, the expectation is that the construct will be removed from future releases of the markup language being defined sometime shortly after the *validUntil* date.

An ODD processor will typically not process a specification element which has a *validUntil* date that is in the past. An ODD processor will typically warn users about constructs which have a *validUntil* date that is in the future. E.g., the documentation for such a construct might include the phrase *warning: deprecated* in red.

---

## att.dimensions provides attributes for describing the size of physical objects.

*Module* tei — The TEI Infrastructure

*Members* att.damaged[damage damageSpan] add addSpan age birth date death del delSpan depth dim dimensions ellipsis ex floruit gap geogFeat height mod offset origDate population redo restore retrace secl space state subst substJoin supplied surplus time trait unclear undo width

*Attributes*
- att.ranging
  - *@atLeast*
  - *@atMost*
  - *@min*
  - *@max*
  - *@confidence*

`@unit` names the unit used for the measurement
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **cm** (centimetres)
**mm** (millimetres)
**in** (inches)
**line** lines of text
**char** (characters) characters of text

`@quantity` specifies the length in the units specified
*Status* Optional
*Datatype* teidata.numeric

`@extent` indicates the size of the object concerned using a project-specific vocabulary combining quantity and units in a single string of words.
*Status* Optional
*Datatype* teidata.text

```
<gap extent="5 words"/>
<height extent="half the page"/>
```

`@precision` characterizes the precision of the values specified by the other attributes.
*Status* Optional
*Datatype* teidata.certainty

`@scope` where the measurement summarizes more than one observation, specifies the applicability of this measurement.
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **all** measurement applies to all instances.
**most** measurement applies to most of the instances inspected.
**range** measurement applies to only the specified range of instances.

---

**att.divLike** provides attributes common to all elements which behave in the same way as divisions. [*4. Default Text Structure*]

*Module* tei — The TEI Infrastructure

*Members* div div1 div2 div3 div4 div5 div6 div7 lg

*Attributes*
- att.fragmentable
  - *@part*
- att.metrical
  - *@met*
  - *@real*

– *@rhyme*

`@org` (organization) specifies how the content of the division is organized.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **composite** no claim is made about the sequence in which the immediate contents of this division are to be processed, or their inter-relationships.

**uniform** the immediate contents of this element are regarded as forming a logical unit, to be processed in sequence.[Default]

`@sample` indicates whether this division is a sample of the original source and if so, from which part.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **initial** division lacks material present at end in source.

**medial** division lacks material at start and end.

**final** division lacks material at start.

**unknown** position of sampled material within original unknown.

**complete** division is not a sample.[Default]

---

## att.docStatus provides attributes for use on metadata elements describing the status of a document.

*Module* tei — The TEI Infrastructure

*Members* bibl biblFull biblStruct change msDesc object post revisionDesc schemaSpec

*Attributes*

`@status` describes the status of a document either currently or, when associated with a dated element, at the time indicated.

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **approved**

**candidate**

**cleared**

**deprecated**

**draft** [Default]

**embargoed**

**expired**

**frozen**

**galley**

**proposed**

**published**

**recommendation**

**submitted**

**unfinished**

**withdrawn**

*Example*

```
<revisionDesc status="published">
  <change when="2010-10-21"
    status="published"/>
```

```
  <change when="2010-10-02" status="cleared"/>
  <change when="2010-08-02"
    status="embargoed"/>
  <change when="2010-05-01" status="frozen"
    who="#MSM"/>
  <change when="2010-03-01" status="draft"
    who="#LB"/>
</revisionDesc>
```

---

**att.duration** provides attributes for normalization of elements that contain datable events.

*Module* spoken — Transcriptions of Speech

*Members* att.timed[annotationBlock binaryObject ellipsis gap incident kinesic media pause post u vocal writing] date recording time

*Attributes*
- att.duration.iso
  - *@dur-iso*
- att.duration.w3c
  - *@dur*

This 'superclass' provides attributes that can be used to provide normalized values of temporal information. By default, the attributes from the `att.duration.w3c` class are provided. If the module for names & dates is loaded, this class also provides attributes from the `att.duration.iso` class. In general, the possible values of attributes restricted to the W3C datatypes form a subset of those values available via the ISO 8601 standard. However, the greater expressiveness of the ISO datatypes is rarely needed, and there exists much greater software support for the W3C datatypes.

---

**att.duration.iso** provides attributes for recording normalized temporal durations. [*3.6.4. Dates and Times 14.4. Dates*]

*Module* tei — The TEI Infrastructure

*Members* att.duration[att.timed[annotationBlock binaryObject ellipsis gap incident kinesic media pause post u vocal writing] date recording time]

*Attributes*

`@dur-iso` (duration) indicates the length of this element in time.
*Status* Optional
*Datatype* teidata.duration.iso

If both *when* and *dur* or *dur-iso* are specified, the values should be interpreted as indicating a span of time by its starting time (or date) and duration. In order to represent a time range by a duration and its ending time the *when-iso* attribute must be used.

In providing a 'regularized' form, no claim is made that the form in the source text is incorrect; the regularized form is simply that chosen as the main form for purposes of unifying variant forms under a single heading.

---

**att.duration.w3c** provides attributes for recording normalized temporal durations. [*3.6.4. Dates and Times 14.4. Dates*]

*Module* tei — The TEI Infrastructure

*Members* att.duration[att.timed[annotationBlock binaryObject ellipsis gap incident kinesic media pause post u vocal writing] date recording time]

*Attributes*

`@dur` (duration) indicates the length of this element in time.

*Status* Optional

*Datatype* teidata.duration.w3c

If both *when* and *dur* are specified, the values should be interpreted as indicating a span of time by its starting time (or date) and duration. In order to represent a time range by a duration and its ending time the *when-iso* attribute must be used.

In providing a 'regularized' form, no claim is made that the form in the source text is incorrect; the regularized form is simply that chosen as the main form for purposes of unifying variant forms under a single heading.

---

## att.edition

**att.edition** provides attributes identifying the source edition from which some encoded feature derives.

*Module* tei — The TEI Infrastructure

*Members* cb gb lb milestone pb refState

*Attributes*

`@ed` (edition) supplies a sigil or other arbitrary identifier for the source edition in which the associated feature (for example, a page, column, or line beginning) occurs at this point in the text.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

`@edRef` (edition reference) provides a pointer to the source edition in which the associated feature (for example, a page, column, or line beginning) occurs at this point in the text.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Example*

```
<l>Of Mans First Disobedience,<lb ed="1674"/> and<lb ed="1667"/> the
Fruit</l>
<l>Of that Forbidden Tree, whose<lb ed="1667 1674"/> mortal tast</l>
<l>Brought Death into the World,<lb ed="1667"/> and all<lb ed="1674"/> our
woe,</l>
```

*Example*

```
<listBibl>
  <bibl xml:id="stapledon1937">
    <author>Olaf Stapledon</author>,
  <title>Starmaker</title>, <publisher>Methuen</publisher>, <date>1937</date>
  </bibl>
  <bibl xml:id="stapledon1968">
    <author>Olaf Stapledon</author>,
  <title>Starmaker</title>, <publisher>Dover</publisher>, <date>1968</date>
  </bibl>
</listBibl>
<!-- ... -->
<p>Looking into the future aeons from the supreme moment of
the cosmos, I saw the populations still with all their
strength maintaining the<pb n="411" edRef="#stapledon1968"/>essentials of
their ancient culture,
still living their personal lives in zest and endless
novelty of action, … I saw myself still
preserving, though with increasing difficulty, my lucid
```

```
con<pb  n="291"  break="no"
   edRef="#stapledon1937"/>sciousness;</p>
```

In the above example, the soft hyphen in Stapledon 1937 is omitted. Such decisions may be documented in the edition's declaration of editorial principles, e.g. with the <hyphenation> element in the <teiHeader>.

*Note* These guidelines provide no semantic basis or suggested precedence when both *ed* and *edRef* are provided. For this reason simultaneous use of both is not recommended unless documentation explaining the use is provided, probably in an ODD customization, for interchange.

---

**att.editLike** provides attributes describing the nature of an encoded scholarly intervention or interpretation of any kind. [*3.5. Simple Editorial Changes 11.3.1. Origination 14.3.2. The Person Element 12.3.1.1. Core Elements for Transcriptional Work*]

*Module* tei — The TEI Infrastructure

*Members* att.transcriptional[add addSpan del delSpan mod redo restore retrace rt subst substJoin undo] affiliation age am birth climate corr date death education event eventName ex expan faith floruit gap gender geogFeat geogName langKnowledge langKnown location name nationality objectName occupation offset org orgName origDate origPlace origin persName persPronouns person persona place placeName population reg relation residence secl sex socecStatus state supplied surplus terrain time trait unclear

*Attributes*

`@evidence` indicates the nature of the evidence supporting the reliability or accuracy of the intervention or interpretation.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace
*Suggested values include:* **internal** there is internal evidence to support the intervention.
**external** there is external evidence to support the intervention.
**conjecture** the intervention or interpretation has been made by the editor, cataloguer, or scholar on the basis of their expertise.

`@instant` indicates whether this is an instant revision or not.
*Status* Optional
*Datatype* teidata.xTruthValue
*Default* false

The members of this attribute class are typically used to represent any kind of editorial intervention in a text, for example a correction or interpretation, or to date or localize manuscripts etc.

Each pointer on the *source* (if present) corresponding to a witness or witness group should reference a bibliographic citation such as a <witness>, <msDesc>, or <bibl> element, or another external bibliographic citation, documenting the source concerned.

---

**att.enjamb** (enjambement) provides attributes that may be used to indicate enjambement of the parent element. [*6.2. Components of the Verse Line*]

*Module* verse — Verse

*Members* l

*Attributes*

`@enjamb` (enjambement) indicates that the end of a verse line is marked by enjambement.
*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **no** the line is end-stopped

**yes** the line in question runs on into the next

**weak** the line is weakly enjambed

**strong** the line is strongly enjambed

*Note* The usual practice will be to give the value 'yes' to this attribute when enjambement is being marked, or the values 'weak' and 'strong' if degrees of enjambement are of interest; if no value is given, however, the attribute does not default to a value of 'no'; this allows the attribute to be omitted entirely when enjambement is not of particular interest.

---

**att.entryLike** provides attributes used to distinguish different styles of dictionary entries. [*10.1. Dictionary Body and Overall Structure 10.2. The Structure of Dictionary Entries*]

*Module* dictionaries — Dictionaries

*Members* entry entryFree superEntry

*Attributes* • att.typed

– ~~type~~

– @subtype

`@type` indicates type of entry, in dictionaries with multiple types.

*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **main** a main entry (default).[Default]

**hom** (homograph) groups information relating to one homograph within an entry.

**xref** (cross reference) a reduced entry whose only function is to point to another main entry (e.g. for forms of an irregular verb or for variant spellings: *was* pointing to *be*, or *esthete* to *aesthete*).

**affix** an entry for a prefix, infix, or suffix.

**abbr** (abbreviation) an entry for an abbreviation.

**supplemental** a supplemental entry (for use in dictionaries which issue supplements to their main work in which they include updated information about entries).

**foreign** an entry for a foreign word in a monolingual dictionary.

The global *n* attribute may be used to encode the homograph numbers attached to entries for homographs.

---

**att.formula** provides attributes for defining a mathematical formula. [*2.3.9. The Unit Declaration*]

*Module* tei — The TEI Infrastructure

*Members* conversion

*Attributes*

`@formula` A *formula* is provided to describe a mathematical calculation such as a conversion between measurement systems.

*Status* Optional

*Datatype* teidata.xpath

*Example*

```
<encodingDesc>
  <unitDecl>
    <unitDef  xml:id="stadium"  type="linear">
      <label>stadium</label>
      <placeName  ref="#rome"/>
      <conversion  fromUnit="#pes"
        toUnit="#stadium"  formula="$fromUnit * 625"/>
      <desc>The stadium was a Roman unit of linear measurement equivalent to 625
pedes, or Roman feet.</desc>
    </unitDef>
  </unitDecl>
</encodingDesc>
```

*Example*

```
<encodingDesc>
  <unitDecl>
    <unitDef  xml:id="wmw"  type="power">
      <label>whatmeworry</label>
      <conversion  fromUnit="#hpk"
        toUnit="#wmw"  formula="$fromUnit * 1"/>
      <desc>In the Potrzebie system of measures as introduced by Donald Knuth,
the whatmeworry unit of power is equivalent to one hah per kovac.</desc>
    </unitDef>
    <unitDef  xml:id="kwmw"  type="power">
      <label>kilowhatmeworry</label>
      <conversion  fromUnit="#wmw"
        toUnit="#kwmw"  formula="$fromUnit div 1000"/>
      <desc>The kilowhatmeworry is equivalent to 1000 whatmeworries.</desc>
    </unitDef>
    <unitDef  xml:id="ap"  type="power">
      <label>kilowhatmeworry</label>
      <conversion  fromUnit="#kwmw"
        toUnit="#ap"  formula="$fromUnit div 100"/>
      <desc>One unit of aeolipower (A.P.) is equivalent to 100
kilowhatmeworries.</desc>
    </unitDef>
  </unitDecl>
</encodingDesc>
```

*Example*

```
<conversion  fromUnit="#furlongsPerFortnight"
  toUnit="#milesPerHour"
  formula="$fromUnit cast as xs:decimal * 0.000372"/>
```

*Example*

```
<conversion  fromUnit="#deciday"
  toUnit="hour"
  formula="$fromUnit cast as xs:decimal * 144 div 60"/>
```

*Note* This attribute class provides *formula* for use in defining a value used in mathematical calculation. It can be used to store a mathematical operation needed to convert from one system of measurement to another. We use the teidata.xpath datatype to express this value in order to communicate mathematical operations on an XML node or nodes. The $fromUnit variable notation simplifies referencing of the

*fromUnit* attribute on the parent <conversion> element. Note that 'div' is required to express the division operator in XPath.

---

**att.fragmentable** provides attributes for representing fragmentation of a structural element, typically as a consequence of some overlapping hierarchy.

*Module* tei — The TEI Infrastructure

*Members* att.divLike[div div1 div2 div3 div4 div5 div6 div7 lg] att.segLike[c cl m pc phr s seg w] ab l p post

*Attributes*

`@part` specifies whether or not its parent element is fragmented in some way, typically by some other overlapping structure: for example a speech which is divided between two or more verse stanzas, a paragraph which is split across a page division, a verse line which is divided between two speakers.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **Y** (yes) the element is fragmented in some (unspecified) respect

**N** (no) the element is not fragmented, or no claim is made as to its completeness[Default]

**I** (initial) this is the initial part of a fragmented element

**M** (medial) this is a medial part of a fragmented element

**F** (final) this is the final part of a fragmented element

*Note* The values I, M, or F should be used only where it is clear how the element may be reconstituted.

---

**att.gaijiProp** provides attributes for defining the properties of non-standard characters or glyphs. [*5. Characters, Glyphs, and Writing Modes*]

*Module* gaiji — Characters, Glyphs, and Writing Modes

*Members* localProp unicodeProp unihanProp

*Attributes*

- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*
    - *@to-custom*
    - *@datingPoint*
    - *@datingMethod*
  - att.datable.iso
    - *@when-iso*
    - *@notBefore-iso*
    - *@notAfter-iso*
    - *@from-iso*
    - *@to-iso*

- att.datable.w3c
  - *@when*
  - *@notBefore*
  - *@notAfter*
  - *@from*
  - *@to*

`@name` provides the name of the character or glyph property being defined.
*Status* Required
*Datatype* teidata.xmlName

`@value` provides the value of the character or glyph property being defined.
*Status* Required
*Datatype* teidata.text

`@version` specifies the version number of the Unicode Standard in which this property name is defined.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **1.0.1**
**1.1**
**2.0**
**2.1**
**3.0**
**3.1**
**3.2**
**4.0**
**4.1**
**5.0**
**5.1**
**5.2**
**6.0**
**6.1**
**6.2**
**6.3**
**7.0**
**8.0**
**9.0**
**10.0**
**11.0**
**12.0**
**12.1**
**13.0**
**14.0**
**15.0**
**unassigned**

`@scheme` supplies the name of the character set system from which this property is drawn.
*Status* Optional
*Datatype* teidata.enumerated

*Sample values include:* **Unicode** (Unicode) ISO 10646[Default]

**Bridwell** (E. Nelson Bridwell) Original character set developed by E. Nelson Bridwell as described by Al Turniansky, in use from the 1950s to 1985.

**Brewer** (Georg Brewer) Developed by Georg Brewer, with a look similar to the Byrne glyphs. Like the Byrne set, this is not a true character set, but rather a set of alternate glyphs.

**Doyle** (Darren Doyle) Glyph set (in some cases associated with multiple characters) developed by Darren Doyle as part of a comprehensive version of the language created in part for an invented language classs at UT Austin in 2006.

**Schreyer** (Christine Schreyer) Character set (without actual codepoints) of 153 characters developed, along with pronunciation rules, roughly 300 words, and a grammar, developed by Christine Schreyer developed in 2012 for Warner Brothers.

*Example* In this example a definition for the Unicode property Decomposition Mapping is provided.

```
<unicodeProp  name="Decomposition_Mapping"
  value="circle"/>
```

*Note* All name-only attributes need an xs:boolean attribute value inside *value*.

---

**att.global** provides attributes common to all elements in the TEI encoding scheme. [*1.3.1.1. Global Attributes*]

*Module* tei — The TEI Infrastructure

*Members* TEI ab abbr abstract accMat acquisition activity actor add addName addSpan additional additions addrLine address adminInfo affiliation age alt altGrp altIdent altIdentifier alternate am analytic anchor annotation annotationBlock anyElement app appInfo application arc argument att attDef attList attRef author authority availability back bibl biblFull biblScope biblStruct bicond binary binaryObject binding bindingDesc birth bloc body broadcast byline c cRefPattern caesura calendar calendarDesc camera caption case castGroup castItem castList catDesc catRef catchwords category cb cell certainty change channel char charDecl choice cit citeData citeStructure citedRange cl classCode classDecl classRef classSpec classes climate closer code collation collection colloc colophon cond condition constitution constraint constraintDecl constraintSpec content conversion corr correction correspAction correspContext correspDesc country creation custEvent custodialHist damage damageSpan dataFacet dataRef dataSpec datatype date dateline death decoDesc decoNote def default defaultVal del delSpan depth derivation desc dictScrap dim dimensions distinct distributor district div div1 div2 div3 div4 div5 div6 div7 divGen docAuthor docDate docEdition docImprint docTitle domain eLeaf eTree edition editionStmt editor editorialDecl education eg egXML elementRef elementSpec ellipsis email emph empty encodingDesc entry entryFree epigraph epilogue equipment equiv etym event eventName ex exemplum expan explicit extent f fDecl fDescr fLib facsimile factuality faith figDesc figure fileDesc filiation finalRubric floatingText floruit foliation foreign forename forest form formula front fs fsConstraints fsDecl fsDescr fsdDecl fsdLink funder fvLib fw g gap gb gen genName gender geo geoDecl geogFeat geogName gi gloss glyph gram gramGrp graph graphic group handDesc handNote handNotes handShift head headItem headLabel height heraldry hi history hom hyph hyphenation iNode iType ident idno if iff imprimatur imprint incident incipit index institution interaction interp interpGrp interpretation item join joinGrp keywords kinesic l label lacunaEnd lacunaStart lang langKnowledge langKnown langUsage language layout layoutDesc lb lbl leaf lem lg licence line link linkGrp list listAnnotation listApp listBibl listChange listEvent listForest listNym listObject listOrg listPerson listPlace listPrefixDef listRef listRelation listTranspose listWit localProp locale location locus locusGrp m macroRef macroSpec mapping material measure measureGrp media meeting memberOf mentioned metDecl metSym metamark milestone mod model modelGrp modelSequence moduleRef moduleSpec monogr mood move msContents msDesc msFrag msIdentifier msItem msItemStruct msName msPart musicNotation name nameLink namespace nationality node normalization notatedMusic note noteGrp notesStmt num

number numeric nym oRef object objectDesc objectIdentifier objectName objectType occupation offset opener org orgName orig origDate origPlace origin orth outputRendition p pRef param paramList paramSpec particDesc path pause pb pc per performance persName persPronouns person personGrp persona phr physDesc place placeName population pos post postBox postCode postscript precision prefixDef preparedness principal profileDesc projectDesc prologue pron provenance ptr pubPlace publicationStmt publisher punctuation purpose q quotation quote rb rdg rdgGrp re recordHist recording recordingStmt redo ref refState refsDecl reg region relatedItem relation remarks rendition repository residence resp respStmt respons restore retrace revisionDesc rhyme role roleDesc roleName root row rs rt rubric ruby s said salute samplingDecl schemaRef schemaSpec scriptDesc scriptNote scriptStmt seal sealDesc secFol secl seg segmentation sense sequence series seriesStmt set setting settingDesc settlement sex shift sic signatures signed soCalled socecStatus sound source sourceDesc sourceDoc sp spGrp space span spanGrp speaker specDesc specGrp specGrpRef specList sponsor stage stamp standOff state stdVals street stress string styleDefDecl subc subst substJoin summary superEntry supplied support supportDesc surface surfaceGrp surname surplus surrogates syll symbol table tag tagUsage tagsDecl taxonomy tech teiCorpus teiHeader term terrain text textClass textDesc textLang textNode then time timeline title titlePage titlePart titleStmt tns trailer trait transcriptionDesc transpose tree triangle typeDesc typeNote u unclear undo unicodeProp unihanProp unit unitDecl unitDef usg vAlt vColl vDefault vLabel vMerge vNot vRange val valDesc valItem valList variantEncoding view vocal w watermark when width wit witDetail witEnd witStart witness writing xenoData xr zone

*Attributes*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@xml:id` (identifier) provides a unique identifier for the element bearing the attribute.

*Status* Optional

*Datatype* ID

*Note* The *xml:id* attribute may be used to specify a canonical reference for an element; see section *3.11. Reference Systems*.

`@n` (number) gives a number (or other label) for an element, which is not necessarily unique within the document.
*Status* Optional
*Datatype* teidata.text
*Note* The value of this attribute is always understood to be a single token, even if it contains space or other punctuation characters, and need not be composed of numbers only. It is typically used to specify the numbering of chapters, sections, list items, etc.; it may also be used in the specification of a standard reference system for the text.

`@xml:lang` (language) indicates the language of the element content using a 'tag' generated according to BCP 47.
*Status* Optional
*Datatype* teidata.language

```
<p> … The consequences of
this rapid depopulation were the loss of the last
<foreign  xml:lang="rap">ariki</foreign> or chief
(Routledge 1920:205,210) and their connections to
ancestral territorial organization.</p>
```

*Note* The *xml:lang* value will be inherited from the immediately enclosing element, or from its parent, and so on up the document hierarchy. It is generally good practice to specify *xml:lang* at the highest appropriate level, noticing that a different default may be needed for the <teiHeader> from that needed for the associated resource element or elements, and that a single TEI document may contain texts in many languages.

Only attributes with free text values (rare in these guidelines) will be in the scope of *xml:lang*.

The authoritative list of registered language subtags is maintained by IANA and is available at https://www.iana.org/assignments/language-subtag-registry. For a good general overview of the construction of language tags, see https://www.w3.org/International/articles/language-tags/, and for a practical step-by-step guide, see https://www.w3.org/International/questions/qa-choosing-language-tags.en.php.

The value used must conform with BCP 47. If the value is a private use code (i.e., starts with x- or contains -x-), a <language> element with a matching value for its *ident* attribute should be supplied in the TEI header to document this value. Such documentation may also optionally be supplied for non-private-use codes, though these must remain consistent with their ( IETF) Internet Engineering Task Force definitions.

`@xml:base` provides a base URI reference with which applications can resolve relative URI references into absolute URI references.
*Status* Optional
*Datatype* teidata.pointer

```
<div  type="bibl">
  <head>Selections from <title  level="m">The Collected Letters
of Robert Southey. Part 1: 1791-1797</title>
  </head>
  <listBibl  xml:base="https://romantic-circles.org/sites/default/files/imported/edit
    <bibl>
      <ref  target="letterEEd.26.3.xml">
        <title>Robert Southey to Grosvenor Charles
Bedford</title>, <date  when="1792-04-03">3 April 1792</date>.
      </ref>
```

```
    </bibl>
    <bibl>
      <ref  target="letterEEd.26.57.xml">
        <title>Robert Southey to Anna Seward</title>,
<date  when="1793-09-18">18 September 1793</date>.
      </ref>
    </bibl>
    <bibl>
      <ref  target="letterEEd.26.85.xml">
        <title>Robert Southey to Robert Lovell</title>,
<date  from="1794-04-05"
        to="1794-04-06">5-6 April, 1794</date>.
      </ref>
    </bibl>
  </listBibl>
</div>
```

`@xml:space` signals an intention about how white space should be managed by applications.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **default** signals that the application's default white-space processing modes are acceptable

**preserve** indicates the intent that applications preserve all white space

*Note* The XML specification provides further guidance on the use of this attribute. Note that many parsers may not handle xml:space correctly.

---

## att.global.analytic provides additional global attributes for associating specific analyses or interpretations with appropriate portions of a text. [*18.2. Global Attributes for Simple Analyses 18.3. Spans and Interpretations*]

*Module* analysis — Simple Analytic Mechanisms

*Members* att.global[TEI ab abbr abstract accMat acquisition activity actor add addName addSpan additional additions addrLine address adminInfo affiliation age alt altGrp altIdent altIdentifier alternate am analytic anchor annotation annotationBlock anyElement app appInfo application arc argument att attDef attList attRef author authority availability back bibl biblFull biblScope biblStruct bicond binary binaryObject binding bindingDesc birth bloc body broadcast byline c cRefPattern caesura calendar calendarDesc camera caption case castGroup castItem castList catDesc catRef catchwords category cb cell certainty change channel char charDecl choice cit citeData citeStructure citedRange cl classCode classDecl classRef classSpec classes climate closer code collation collection colloc colophon cond condition constitution constraint constraintDecl constraintSpec content conversion corr correction correspAction correspContext correspDesc country creation custEvent custodialHist damage damageSpan dataFacet dataRef dataSpec datatype date dateline death decoDesc decoNote def default defaultVal del delSpan depth derivation desc dictScrap dim dimensions distinct distributor district div div1 div2 div3 div4 div5 div6 div7 divGen docAuthor docDate docEdition docImprint docTitle domain eLeaf eTree edition editionStmt editor editorialDecl education eg egXML elementRef elementSpec ellipsis email emph empty encodingDesc entry entryFree epigraph epilogue equipment equiv etym event eventName ex exemplum expan explicit extent f fDecl fDescr fLib facsimile factuality faith figDesc figure fileDesc filiation finalRubric floatingText floruit foliation foreign forename forest form formula front fs fsConstraints fsDecl fsDescr fsdDecl fsdLink funder fvLib fw g gap gb gen genName gender geo geoDecl geogFeat geogName gi gloss glyph gram gramGrp graph graphic group handDesc handNote handNotes handShift head headItem headLabel height heraldry hi history hom hyph hyphenation iNode iType ident idno if iff imprimatur imprint incident incipit index institution interaction interp interpGrp interpretation item join joinGrp keywords kinesic l label lacunaEnd lacunaStart lang langKnowledge langKnown langUsage language layout layoutDesc lb lbl leaf lem lg licence line link linkGrp list listAnnotation listApp listBibl

listChange listEvent listForest listNym listObject listOrg listPerson listPlace listPrefixDef listRef listRelation listTranspose listWit localProp locale location locus locusGrp m macroRef macroSpec mapping material measure measureGrp media meeting memberOf mentioned metDecl metSym metamark milestone mod model modelGrp modelSequence moduleRef moduleSpec monogr mood move msContents msDesc msFrag msIdentifier msItem msItemStruct msName msPart musicNotation name nameLink namespace nationality node normalization notatedMusic note noteGrp notesStmt num number numeric nym oRef object objectDesc objectIdentifier objectName objectType occupation offset opener org orgName orig origDate origPlace origin orth outputRendition p pRef param paramList paramSpec particDesc path pause pb pc per performance persName persPronouns person personGrp persona phr physDesc place placeName population pos post postBox postCode postscript precision prefixDef preparedness principal profileDesc projectDesc prologue pron provenance ptr pubPlace publicationStmt publisher punctuation purpose q quotation quote rb rdg rdgGrp re recordHist recording recordingStmt redo ref refState refsDecl reg region relatedItem relation remarks rendition repository residence resp respStmt respons restore retrace revisionDesc rhyme role roleDesc roleName root row rs rt rubric ruby s said salute samplingDecl schemaRef schemaSpec scriptDesc scriptNote scriptStmt seal sealDesc secFol secl seg segmentation sense sequence series seriesStmt set setting settingDesc settlement sex shift sic signatures signed soCalled socecStatus sound source sourceDesc sourceDoc sp spGrp space span spanGrp speaker specDesc specGrp specGrpRef specList sponsor stage stamp standOff state stdVals street stress string styleDefDecl subc subst substJoin summary superEntry supplied support supportDesc surface surfaceGrp surname surplus surrogates syll symbol table tag tagUsage tagsDecl taxonomy tech teiCorpus teiHeader term terrain text textClass textDesc textLang textNode then time timeline title titlePage titlePart titleStmt tns trailer trait transcriptionDesc transpose tree triangle typeDesc typeNote u unclear undo unicodeProp unihanProp unit unitDecl unitDef usg vAlt vColl vDefault vLabel vMerge vNot vRange val valDesc valItem valList variantEncoding view vocal w watermark when width wit witDetail witEnd witStart witness writing xenoData xr zone]

*Attributes*

`@ana` (analysis) indicates one or more elements containing interpretations of the element on which the *ana* attribute appears.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Note* When multiple values are given, they may reflect either multiple divergent interpretations of an ambiguous text, or multiple mutually consistent interpretations of the same passage in different contexts.

---

**att.global.change** provides attributes allowing its member elements to specify one or more states or revision campaigns with which they are associated.

*Module* transcr — Representation of Primary Sources

*Members* att.global[TEI ab abbr abstract accMat acquisition activity actor add addName addSpan additional additions addrLine address adminInfo affiliation age alt altGrp altIdent altIdentifier alternate am analytic anchor annotation annotationBlock anyElement app appInfo application arc argument att attDef attList attRef author authority availability back bibl biblFull biblScope biblStruct bicond binary binaryObject binding bindingDesc birth bloc body broadcast byline c cRefPattern caesura calendar calendarDesc camera caption case castGroup castItem castList catDesc catRef catchwords category cb cell certainty change channel char charDecl choice cit citeData citeStructure citedRange cl classCode classDecl classRef classSpec classes climate closer code collation collection colloc colophon cond condition constitution constraint constraintDecl constraintSpec content conversion corr correction correspAction correspContext correspDesc country creation custEvent custodialHist damage damageSpan dataFacet dataRef dataSpec datatype date dateline death decoDesc decoNote def default defaultVal del delSpan depth derivation desc dictScrap dim dimensions distinct distributor district div div1 div2 div3 div4 div5 div6 div7 divGen docAuthor docDate docEdition docImprint docTitle domain eLeaf eTree edition editionStmt editor editorialDecl education eg egXML elementRef elementSpec ellipsis email emph empty

encodingDesc entry entryFree epigraph epilogue equipment equiv etym event eventName ex exemplum expan explicit extent f fDecl fDescr fLib facsimile factuality faith figDesc figure fileDesc filiation finalRubric floatingText floruit foliation foreign forename forest form formula front fs fsConstraints fsDecl fsDescr fsdDecl fsdLink funder fvLib fw g gap gb gen genName gender geo geoDecl geogFeat geogName gi gloss glyph gram gramGrp graph graphic group handDesc handNote handNotes handShift head headItem headLabel height heraldry hi history hom hyph hyphenation iNode iType ident idno if iff imprimatur imprint incident incipit index institution interaction interp interpGrp interpretation item join joinGrp keywords kinesic l label lacunaEnd lacunaStart lang langKnowledge langKnown langUsage language layout layoutDesc lb lbl leaf lem lg licence line link linkGrp list listAnnotation listApp listBibl listChange listEvent listForest listNym listObject listOrg listPerson listPlace listPrefixDef listRef listRelation listTranspose listWit localProp locale location locus locusGrp m macroRef macroSpec mapping material measure measureGrp media meeting memberOf mentioned metDecl metSym metamark milestone mod model modelGrp modelSequence moduleRef moduleSpec monogr mood move msContents msDesc msFrag msIdentifier msItem msItemStruct msName msPart musicNotation name nameLink namespace nationality node normalization notatedMusic note noteGrp notesStmt num number numeric nym oRef object objectDesc objectIdentifier objectName objectType occupation offset opener org orgName orig origDate origPlace origin orth outputRendition p pRef param paramList paramSpec particDesc path pause pb pc per performance persName persPronouns person personGrp persona phr physDesc place placeName population pos post postBox postCode postscript precision prefixDef preparedness principal profileDesc projectDesc prologue pron provenance ptr pubPlace publicationStmt publisher punctuation purpose q quotation quote rb rdg rdgGrp re recordHist recording recordingStmt redo ref refState refsDecl reg region relatedItem relation remarks rendition repository residence resp respStmt respons restore retrace revisionDesc rhyme role roleDesc roleName root row rs rt rubric ruby s said salute samplingDecl schemaRef schemaSpec scriptDesc scriptNote scriptStmt seal sealDesc secFol secl seg segmentation sense sequence series seriesStmt set setting settingDesc settlement sex shift sic signatures signed soCalled socecStatus sound source sourceDesc sourceDoc sp spGrp space span spanGrp speaker specDesc specGrp specGrpRef specList sponsor stage stamp standOff state stdVals street stress string styleDefDecl subc subst substJoin summary superEntry supplied support supportDesc surface surfaceGrp surname surplus surrogates syll symbol table tag tagUsage tagsDecl taxonomy tech teiCorpus teiHeader term terrain text textClass textDesc textLang textNode then time timeline title titlePage titlePart titleStmt tns trailer trait transcriptionDesc transpose tree triangle typeDesc typeNote u unclear undo unicodeProp unihanProp unit unitDecl unitDef usg vAlt vColl vDefault vLabel vMerge vNot vRange val valDesc valItem valList variantEncoding view vocal w watermark when width wit witDetail witEnd witStart witness writing xenoData xr zone]

*Attributes*

`@change` points to one or more <change> elements documenting a state or revision campaign to which the element bearing this attribute and its children have been assigned by the encoder.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

---

**att.global.facs** provides attributes used to express correspondence between an element and all or part of a facsimile image or surface. [*12.1. Digital Facsimiles*]

*Module* transcr — Representation of Primary Sources

*Members* att.global[TEI ab abbr abstract accMat acquisition activity actor add addName addSpan additional additions addrLine address adminInfo affiliation age alt altGrp altIdent altIdentifier alternate am analytic anchor annotation annotationBlock anyElement app appInfo application arc argument att attDef attList attRef author authority availability back bibl biblFull biblScope biblStruct bicond binary binaryObject binding bindingDesc birth bloc body broadcast byline c cRefPattern caesura calendar calendarDesc camera caption case castGroup castItem castList catDesc catRef catchwords category cb cell certainty change channel char charDecl choice cit citeData citeStructure citedRange cl classCode classDecl classRef classSpec classes climate closer code collation collection colloc colophon cond condition constitution

constraint constraintDecl constraintSpec content conversion corr correction correspAction correspContext correspDesc country creation custEvent custodialHist damage damageSpan dataFacet dataRef dataSpec datatype date dateline death decoDesc decoNote def default defaultVal del delSpan depth derivation desc dictScrap dim dimensions distinct distributor district div div1 div2 div3 div4 div5 div6 div7 divGen docAuthor docDate docEdition docImprint docTitle domain eLeaf eTree edition editionStmt editor editorialDecl education eg egXML elementRef elementSpec ellipsis email emph empty encodingDesc entry entryFree epigraph epilogue equipment equiv etym event eventName ex exemplum expan explicit extent f fDecl fDescr fLib facsimile factuality faith figDesc figure fileDesc filiation finalRubric floatingText floruit foliation foreign forename forest form formula front fs fsConstraints fsDecl fsDescr fsdDecl fsdLink funder fvLib fw g gap gb gen genName gender geo geoDecl geogFeat geogName gi gloss glyph gram gramGrp graph graphic group handDesc handNote handNotes handShift head headItem headLabel height heraldry hi history hom hyph hyphenation iNode iType ident idno if iff imprimatur imprint incident incipit index institution interaction interp interpGrp interpretation item join joinGrp keywords kinesic l label lacunaEnd lacunaStart lang langKnowledge langKnown langUsage language layout layoutDesc lb lbl leaf lem lg licence line link linkGrp list listAnnotation listApp listBibl listChange listEvent listForest listNym listObject listOrg listPerson listPlace listPrefixDef listRef listRelation listTranspose listWit localProp locale location locus locusGrp m macroRef macroSpec mapping material measure measureGrp media meeting memberOf mentioned metDecl metSym metamark milestone mod model modelGrp modelSequence moduleRef moduleSpec monogr mood move msContents msDesc msFrag msIdentifier msItem msItemStruct msName msPart musicNotation name nameLink namespace nationality node normalization notatedMusic note noteGrp notesStmt num number numeric nym oRef object objectDesc objectIdentifier objectName objectType occupation offset opener org orgName orig origDate origPlace origin orth outputRendition p pRef param paramList paramSpec particDesc path pause pb pc per performance persName persPronouns person personGrp persona phr physDesc place placeName population pos post postBox postCode postscript precision prefixDef preparedness principal profileDesc projectDesc prologue pron provenance ptr pubPlace publicationStmt publisher punctuation purpose q quotation quote rb rdg rdgGrp re recordHist recording recordingStmt redo ref refState refsDecl reg region relatedItem relation remarks rendition repository residence resp respStmt respons restore retrace revisionDesc rhyme role roleDesc roleName root row rs rt rubric ruby s said salute samplingDecl schemaRef schemaSpec scriptDesc scriptNote scriptStmt seal sealDesc secFol secl seg segmentation sense sequence series seriesStmt set setting settingDesc settlement sex shift sic signatures signed soCalled socecStatus sound source sourceDesc sourceDoc sp spGrp space span spanGrp speaker specDesc specGrp specGrpRef specList sponsor stage stamp standOff state stdVals street stress string styleDefDecl subc subst substJoin summary superEntry supplied support supportDesc surface surfaceGrp surname surplus surrogates syll symbol table tag tagUsage tagsDecl taxonomy tech teiCorpus teiHeader term terrain text textClass textDesc textLang textNode then time timeline title titlePage titlePart titleStmt tns trailer trait transcriptionDesc transpose tree triangle typeDesc typeNote u unclear undo unicodeProp unihanProp unit unitDecl unitDef usg vAlt vColl vDefault vLabel vMerge vNot vRange val valDesc valItem valList variantEncoding view vocal w watermark when width wit witDetail witEnd witStart witness writing xenoData xr zone]

*Attributes*

`@facs` (facsimile) points to one or more images, portions of an image, or surfaces which correspond to the current element.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

---

## att.global.linking provides a set of attributes for hypertextual linking. [*17. Linking, Segmentation, and Alignment*]

*Module* linking — Linking, Segmentation, and Alignment

*Members* att.global[TEI ab abbr abstract accMat acquisition activity actor add addName addSpan additional additions addrLine address adminInfo affiliation age alt altGrp altIdent altIdentifier alternate am analytic

anchor annotation annotationBlock anyElement app appInfo application arc argument att attDef attList attRef author authority availability back bibl biblFull biblScope biblStruct bicond binary binaryObject binding bindingDesc birth bloc body broadcast byline c cRefPattern caesura calendar calendarDesc camera caption case castGroup castItem castList catDesc catRef catchwords category cb cell certainty change channel char charDecl choice cit citeData citeStructure citedRange cl classCode classDecl classRef classSpec classes climate closer code collation collection colloc colophon cond condition constitution constraint constraintDecl constraintSpec content conversion corr correction correspAction correspContext correspDesc country creation custEvent custodialHist damage damageSpan dataFacet dataRef dataSpec datatype date dateline death decoDesc decoNote def default defaultVal del delSpan depth derivation desc dictScrap dim dimensions distinct distributor district div div1 div2 div3 div4 div5 div6 div7 divGen docAuthor docDate docEdition docImprint docTitle domain eLeaf eTree edition editionStmt editor editorialDecl education eg egXML elementRef elementSpec ellipsis email emph empty encodingDesc entry entryFree epigraph epilogue equipment equiv etym event eventName ex exemplum expan explicit extent f fDecl fDescr fLib facsimile factuality faith figDesc figure fileDesc filiation finalRubric floatingText floruit foliation foreign forename forest form formula front fs fsConstraints fsDecl fsDescr fsdDecl fsdLink funder fvLib fw g gap gb gen genName gender geo geoDecl geogFeat geogName gi gloss glyph gram gramGrp graph graphic group handDesc handNote handNotes handShift head headItem headLabel height heraldry hi history hom hyph hyphenation iNode iType ident idno if iff imprimatur imprint incident incipit index institution interaction interp interpGrp interpretation item join joinGrp keywords kinesic l label lacunaEnd lacunaStart lang langKnowledge langKnown langUsage language layout layoutDesc lb lbl leaf lem lg licence line link linkGrp list listAnnotation listApp listBibl listChange listEvent listForest listNym listObject listOrg listPerson listPlace listPrefixDef listRef listRelation listTranspose listWit localProp locale location locus locusGrp m macroRef macroSpec mapping material measure measureGrp media meeting memberOf mentioned metDecl metSym metamark milestone mod model modelGrp modelSequence moduleRef moduleSpec monogr mood move msContents msDesc msFrag msIdentifier msItem msItemStruct msName msPart musicNotation name nameLink namespace nationality node normalization notatedMusic note noteGrp notesStmt num number numeric nym oRef object objectDesc objectIdentifier objectName objectType occupation offset opener org orgName orig origDate origPlace origin orth outputRendition p pRef param paramList paramSpec particDesc path pause pb pc per performance persName persPronouns person personGrp persona phr physDesc place placeName population pos post postBox postCode postscript precision prefixDef preparedness principal profileDesc projectDesc prologue pron provenance ptr pubPlace publicationStmt publisher punctuation purpose q quotation quote rb rdg rdgGrp re recordHist recording recordingStmt redo ref refState refsDecl reg region relatedItem relation remarks rendition repository residence resp respStmt respons restore retrace revisionDesc rhyme role roleDesc roleName root row rs rt rubric ruby s said salute samplingDecl schemaRef schemaSpec scriptDesc scriptNote scriptStmt seal sealDesc secFol secl seg segmentation sense sequence series seriesStmt set setting settingDesc settlement sex shift sic signatures signed soCalled socecStatus sound source sourceDesc sourceDoc sp spGrp space span spanGrp speaker specDesc specGrp specGrpRef specList sponsor stage stamp standOff state stdVals street stress string styleDefDecl subc subst substJoin summary superEntry supplied support supportDesc surface surfaceGrp surname surplus surrogates syll symbol table tag tagUsage tagsDecl taxonomy tech teiCorpus teiHeader term terrain text textClass textDesc textLang textNode then time timeline title titlePage titlePart titleStmt tns trailer trait transcriptionDesc transpose tree triangle typeDesc typeNote u unclear undo unicodeProp unihanProp unit unitDecl unitDef usg vAlt vColl vDefault vLabel vMerge vNot vRange val valDesc valItem valList variantEncoding view vocal w watermark when width wit witDetail witEnd witStart witness writing xenoData xr zone]

*Attributes*

`@corresp` (corresponds) points to elements that correspond to the current element in some way.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

```
<group>
  <text  xml:id="t1-g1-t1"
    xml:lang="mi">
    <body  xml:id="t1-g1-t1-body1">
```

```
      <div  type="chapter">
        <head>He Whakamaramatanga mo te Ture Hoko, Riihi hoki, i
nga Whenua Maori, 1876.</head>
        <p>…</p>
      </div>
    </body>
  </text>
  <text  xml:id="t1-g1-t2"
    xml:lang="en">
    <body  xml:id="t1-g1-t2-body1"
      corresp="#t1-g1-t1-body1">
      <div  type="chapter">
        <head>An Act to regulate the Sale, Letting, and Disposal
of Native Lands, 1876.</head>
        <p>…</p>
      </div>
    </body>
  </text>
</group>In this example a <group> contains two <text>s, each
containing the same document in a different language. The
correspondence is indicated using corresp. The language is
indicated using xml:lang, whose value is inherited; both the tag
with the corresp and the tag pointed to by the corresp inherit
the value from their immediate parent.

<!-- In a placeography called "places.xml"
--><place  xml:id="LOND1"
  corresp="people.xml#LOND2 people.xml#GENI1">
  <placeName>London</placeName>
  <desc>The city of London...</desc>
</place>
<!-- In a literary personography called "people.xml" -->
<person  xml:id="LOND2"
  corresp="places.xml#LOND1 #GENI1">
  <persName  type="lit">London</persName>
  <note>
    <p>Allegorical character representing the city of
<placeName  ref="places.xml#LOND1">London</placeName>.</p>
  </note>
</person>
<person  xml:id="GENI1"
  corresp="places.xml#LOND1 #LOND2">
  <persName  type="lit">London’s Genius</persName>
  <note>
    <p>Personification of London’s genius. Appears as an
        allegorical character in mayoral shows.
    </p>
  </note>
</person>In this example, a <place> element containing
information about the city of London is linked with two <person>
elements in a literary personography. This correspondence
represents a slightly looser relationship than the one in the
preceding example; there is no sense in which an allegorical
character could be substituted for the physical city, or vice
versa, but there is obviously a correspondence between them.
```

`@synch` (synchronous) points to elements that are synchronous with the current element.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

`@sameAs` points to an element that is the same as the current element.
*Status* Optional
*Datatype* teidata.pointer

`@copyOf` points to an element of which the current element is a copy.
*Status* Optional
*Datatype* teidata.pointer
*Note* Any content of the current element should be ignored. Its true content is that of the element being pointed at.

`@next` (next) points to the next element of a virtual aggregate of which the current element is part.
*Status* Optional
*Datatype* teidata.pointer
*Note* It is recommended that the element indicated be of the same type as the element bearing this attribute.

`@prev` (previous) points to the previous element of a virtual aggregate of which the current element is part.
*Status* Optional
*Datatype* teidata.pointer
*Note* It is recommended that the element indicated be of the same type as the element bearing this attribute.

`@exclude` points to elements that are in exclusive alternation with the current element.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

`@select` selects one or more alternants; if one alternant is selected, the ambiguity or uncertainty is marked as resolved. If more than one alternant is selected, the degree of ambiguity or uncertainty is marked as reduced by the number of alternants not selected.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace
*Note* This attribute should be placed on an element which is superordinate to all of the alternants from which the selection is being made.

---

## att.global.rendition provides rendering attributes common to all elements in the TEI encoding scheme.

[*1.3.1.1.3. Rendition Indicators*]

*Module* tei — The TEI Infrastructure

*Members* att.global[TEI ab abbr abstract accMat acquisition activity actor add addName addSpan additional additions addrLine address adminInfo affiliation age alt altGrp altIdent altIdentifier alternate am analytic anchor annotation annotationBlock anyElement app appInfo application arc argument att attDef attList attRef author authority availability back bibl biblFull biblScope biblStruct bicond binary binaryObject binding bindingDesc birth bloc body broadcast byline c cRefPattern caesura calendar calendarDesc camera caption case castGroup castItem castList catDesc catRef catchwords category cb cell certainty change channel char charDecl choice cit citeData citeStructure citedRange cl classCode classDecl classRef classSpec classes climate closer code collation collection colloc colophon cond condition constitution constraint constraintDecl constraintSpec content conversion corr correction correspAction correspContext correspDesc country creation custEvent custodialHist damage damageSpan dataFacet dataRef dataSpec datatype date dateline death decoDesc decoNote def default defaultVal del delSpan depth derivation desc dictScrap dim dimensions distinct distributor district div div1 div2 div3 div4 div5 div6 div7 divGen docAuthor docDate docEdition docImprint docTitle domain eLeaf eTree edition editionStmt editor editorialDecl education eg egXML elementRef elementSpec ellipsis email emph empty encodingDesc entry entryFree epigraph epilogue equipment equiv etym event eventName ex exemplum expan explicit extent f fDecl fDescr fLib facsimile factuality faith figDesc figure fileDesc filiation

finalRubric floatingText floruit foliation foreign forename forest form formula front fs fsConstraints fsDecl fsDescr fsdDecl fsdLink funder fvLib fw g gap gb gen genName gender geo geoDecl geogFeat geogName gi gloss glyph gram gramGrp graph graphic group handDesc handNote handNotes handShift head headItem headLabel height heraldry hi history hom hyph hyphenation iNode iType ident idno if iff imprimatur imprint incident incipit index institution interaction interp interpGrp interpretation item join joinGrp keywords kinesic l label lacunaEnd lacunaStart lang langKnowledge langKnown langUsage language layout layoutDesc lb lbl leaf lem lg licence line link linkGrp list listAnnotation listApp listBibl listChange listEvent listForest listNym listObject listOrg listPerson listPlace listPrefixDef listRef listRelation listTranspose listWit localProp locale location locus locusGrp m macroRef macroSpec mapping material measure measureGrp media meeting memberOf mentioned metDecl metSym metamark milestone mod model modelGrp modelSequence moduleRef moduleSpec monogr mood move msContents msDesc msFrag msIdentifier msItem msItemStruct msName msPart musicNotation name nameLink namespace nationality node normalization notatedMusic note noteGrp notesStmt num number numeric nym oRef object objectDesc objectIdentifier objectName objectType occupation offset opener org orgName orig origDate origPlace origin orth outputRendition p pRef param paramList paramSpec particDesc path pause pb pc per performance persName persPronouns person personGrp persona phr physDesc place placeName population pos post postBox postCode postscript precision prefixDef preparedness principal profileDesc projectDesc prologue pron provenance ptr pubPlace publicationStmt publisher punctuation purpose q quotation quote rb rdg rdgGrp re recordHist recording recordingStmt redo ref refState refsDecl reg region relatedItem relation remarks rendition repository residence resp respStmt respons restore retrace revisionDesc rhyme role roleDesc roleName root row rs rt rubric ruby s said salute samplingDecl schemaRef schemaSpec scriptDesc scriptNote scriptStmt seal sealDesc secFol secl seg segmentation sense sequence series seriesStmt set setting settingDesc settlement sex shift sic signatures signed soCalled socecStatus sound source sourceDesc sourceDoc sp spGrp space span spanGrp speaker specDesc specGrp specGrpRef specList sponsor stage stamp standOff state stdVals street stress string styleDefDecl subc subst substJoin summary superEntry supplied support supportDesc surface surfaceGrp surname surplus surrogates syll symbol table tag tagUsage tagsDecl taxonomy tech teiCorpus teiHeader term terrain text textClass textDesc textLang textNode then time timeline title titlePage titlePart titleStmt tns trailer trait transcriptionDesc transpose tree triangle typeDesc typeNote u unclear undo unicodeProp unihanProp unit unitDecl unitDef usg vAlt vColl vDefault vLabel vMerge vNot vRange val valDesc valItem valList variantEncoding view vocal w watermark when width wit witDetail witEnd witStart witness writing xenoData xr zone]

*Attributes*

`@rend` (rendition) indicates how the element in question was rendered or presented in the source text.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

```
<head  rend="align(center) case(allcaps)">
  <lb/>To The <lb/>Duchesse <lb/>of <lb/>Newcastle,
<lb/>On Her <lb/>
  <hi  rend="case(mixed)">New Blazing-World</hi>.
</head>
```

*Note* These Guidelines make no binding recommendations for the values of the *rend* attribute; the characteristics of visual presentation vary too much from text to text and the decision to record or ignore individual characteristics varies too much from project to project. Some potentially useful conventions are noted from time to time at appropriate points in the Guidelines. The values of the *rend* attribute are a set of sequence-indeterminate individual tokens separated by whitespace.

`@style` contains an expression in some formal style definition language which defines the rendering or presentation used for this element in the source text.

*Status* Optional

*Datatype* teidata.text

```
<head  style="text-align: center; font-variant: small-caps">
  <lb/>To The <lb/>Duchesse <lb/>of <lb/>Newcastle, <lb/>On Her
<lb/>
  <hi  style="font-variant: normal">New Blazing-World</hi>.
</head>
```

*Note* Unlike the attribute values of *rend*, which uses whitespace as a separator, the *style* attribute may contain whitespace. This attribute is intended for recording inline stylistic information concerning the source, not any particular output.

The formal language in which values for this attribute are expressed may be specified using the <styleDefDecl> element in the TEI header.

If *style* and *rendition* are both present on an element, then *style* overrides or complements *rendition*. *style* should not be used in conjunction with *rend*, because the latter does not employ a formal style definition language.

`@rendition` points to a description of the rendering or presentation used for this element in the source text.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

```
<head  rendition="#ac #sc">
  <lb/>To The <lb/>Duchesse <lb/>of <lb/>Newcastle, <lb/>On Her
<lb/>
  <hi  rendition="#normal">New Blazing-World</hi>.
</head>
<!-- elsewhere... -->
<rendition  xml:id="sc"
  scheme="css">font-variant: small-caps</rendition>
<rendition  xml:id="normal"
  scheme="css">font-variant: normal</rendition>
<rendition  xml:id="ac"
  scheme="css">text-align: center</rendition>
```

*Note* The *rendition* attribute is used in a very similar way to the *class* attribute defined for XHTML but with the important distinction that its function is to describe the appearance of the source text, not necessarily to determine how that text should be presented on screen or paper.

If *rendition* is used to refer to a style definition in a formal language like CSS, it is recommended that it not be used in conjunction with *rend*. Where both *rendition* and *rend* are supplied, the latter is understood to override or complement the former.

Each URI provided should indicate a <rendition> element defining the intended rendition in terms of some appropriate style language, as indicated by the *scheme* attribute.

---

## att.global.responsibility provides attributes indicating the agent responsible for some aspect of the text, the markup or something asserted by the markup, and the degree of certainty associated with it.

[*1.3.1.1.4. Sources, certainty, and responsibility 3.5. Simple Editorial Changes 12.3.2.2. Hand, Responsibility, and Certainty Attributes 18.3. Spans and Interpretations 14.1.1. Linking Names and Their Referents*]

*Module* tei — The TEI Infrastructure

*Members* att.global[TEI ab abbr abstract accMat acquisition activity actor add addName addSpan additional

additions addrLine address adminInfo affiliation age alt altGrp altIdent altIdentifier alternate am analytic anchor annotation annotationBlock anyElement app appInfo application arc argument att attDef attList attRef author authority availability back bibl biblFull biblScope biblStruct bicond binary binaryObject binding bindingDesc birth bloc body broadcast byline c cRefPattern caesura calendar calendarDesc camera caption case castGroup castItem castList catDesc catRef catchwords category cb cell certainty change channel char charDecl choice cit citeData citeStructure citedRange cl classCode classDecl classRef classSpec classes climate closer code collation collection colloc colophon cond condition constitution constraint constraintDecl constraintSpec content conversion corr correction correspAction correspContext correspDesc country creation custEvent custodialHist damage damageSpan dataFacet dataRef dataSpec datatype date dateline death decoDesc decoNote def default defaultVal del delSpan depth derivation desc dictScrap dim dimensions distinct distributor district div div1 div2 div3 div4 div5 div6 div7 divGen docAuthor docDate docEdition docImprint docTitle domain eLeaf eTree edition editionStmt editor editorialDecl education eg egXML elementRef elementSpec ellipsis email emph empty encodingDesc entry entryFree epigraph epilogue equipment equiv etym event eventName ex exemplum expan explicit extent f fDecl fDescr fLib facsimile factuality faith figDesc figure fileDesc filiation finalRubric floatingText floruit foliation foreign forename forest form formula front fs fsConstraints fsDecl fsDescr fsdDecl fsdLink funder fvLib fw g gap gb gen genName gender geo geoDecl geogFeat geogName gi gloss glyph gram gramGrp graph graphic group handDesc handNote handNotes handShift head headItem headLabel height heraldry hi history hom hyph hyphenation iNode iType ident idno if iff imprimatur imprint incident incipit index institution interaction interp interpGrp interpretation item join joinGrp keywords kinesic l label lacunaEnd lacunaStart lang langKnowledge langKnown langUsage language layout layoutDesc lb lbl leaf lem lg licence line link linkGrp list listAnnotation listApp listBibl listChange listEvent listForest listNym listObject listOrg listPerson listPlace listPrefixDef listRef listRelation listTranspose listWit localProp locale location locus locusGrp m macroRef macroSpec mapping material measure measureGrp media meeting memberOf mentioned metDecl metSym metamark milestone mod model modelGrp modelSequence moduleRef moduleSpec monogr mood move msContents msDesc msFrag msIdentifier msItem msItemStruct msName msPart musicNotation name nameLink namespace nationality node normalization notatedMusic note noteGrp notesStmt num number numeric nym oRef object objectDesc objectIdentifier objectName objectType occupation offset opener org orgName orig origDate origPlace origin orth outputRendition p pRef param paramList paramSpec particDesc path pause pb pc per performance persName persPronouns person personGrp persona phr physDesc place placeName population pos post postBox postCode postscript precision prefixDef preparedness principal profileDesc projectDesc prologue pron provenance ptr pubPlace publicationStmt publisher punctuation purpose q quotation quote rb rdg rdgGrp re recordHist recording recordingStmt redo ref refState refsDecl reg region relatedItem relation remarks rendition repository residence resp respStmt respons restore retrace revisionDesc rhyme role roleDesc roleName root row rs rt rubric ruby s said salute samplingDecl schemaRef schemaSpec scriptDesc scriptNote scriptStmt seal sealDesc secFol secl seg segmentation sense sequence series seriesStmt set setting settingDesc settlement sex shift sic signatures signed soCalled socecStatus sound source sourceDesc sourceDoc sp spGrp space span spanGrp speaker specDesc specGrp specGrpRef specList sponsor stage stamp standOff state stdVals street stress string styleDefDecl subc subst substJoin summary superEntry supplied support supportDesc surface surfaceGrp surname surplus surrogates syll symbol table tag tagUsage tagsDecl taxonomy tech teiCorpus teiHeader term terrain text textClass textDesc textLang textNode then time timeline title titlePage titlePart titleStmt tns trailer trait transcriptionDesc transpose tree triangle typeDesc typeNote u unclear undo unicodeProp unihanProp unit unitDecl unitDef usg vAlt vColl vDefault vLabel vMerge vNot vRange val valDesc valItem valList variantEncoding view vocal w watermark when width wit witDetail witEnd witStart witness writing xenoData xr zone]

*Attributes*

`@cert` (certainty) signifies the degree of certainty associated with the intervention or interpretation.
*Status* Optional
*Datatype* teidata.probCert

`@resp` (responsible party) indicates the agency responsible for the intervention or interpretation, for example an editor or transcriber.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Note* To reduce the ambiguity of a *resp* pointing directly to a person or organization, we recommend that *resp* be used to point not to an agent (<person> or <org>) but to a <respStmt>, <author>, <editor> or similar element which clarifies the exact role played by the agent. Pointing to multiple <respStmt>s allows the encoder to specify clearly each of the roles played in part of a TEI file (creating, transcribing, encoding, editing, proofing etc.).

*Example*

```
Blessed are the
<choice>
  <sic>cheesemakers</sic>
  <corr  resp="#editor"  cert="high">peacemakers</corr>
</choice>: for they shall be called the children of God.
```

*Example*

```
<!-- in the <text> ... --><lg>
<!-- ... -->
  <l>Punkes, Panders, baſe extortionizing
     sla<choice>
      <sic>n</sic>
      <corr  resp="#JENS1_transcriber">u</corr>
    </choice>es,</l>
<!-- ... -->
</lg>
<!-- in the <teiHeader> ... -->
<!-- ... -->
<respStmt  xml:id="JENS1_transcriber">
  <resp  when="2014">Transcriber</resp>
  <name>Janelle Jenstad</name>
</respStmt>
```

---

## att.global.source provides attributes used by elements to point to an external source. [*1.3.1.1.4. Sources, certainty, and responsibility 3.3.3. Quotation 8.3.4. Writing*]

*Module* tei — The TEI Infrastructure

*Members* att.global[TEI ab abbr abstract accMat acquisition activity actor add addName addSpan additional additions addrLine address adminInfo affiliation age alt altGrp altIdent altIdentifier alternate am analytic anchor annotation annotationBlock anyElement app appInfo application arc argument att attDef attList attRef author authority availability back bibl biblFull biblScope biblStruct bicond binary binaryObject binding bindingDesc birth bloc body broadcast byline c cRefPattern caesura calendar calendarDesc camera caption case castGroup castItem castList catDesc catRef catchwords category cb cell certainty change channel char charDecl choice cit citeData citeStructure citedRange cl classCode classDecl classRef classSpec classes climate closer code collation collection colloc colophon cond condition constitution constraint constraintDecl constraintSpec content conversion corr correction correspAction correspContext correspDesc country creation custEvent custodialHist damage damageSpan dataFacet dataRef dataSpec datatype date dateline death decoDesc decoNote def default defaultVal del delSpan depth derivation desc dictScrap dim dimensions distinct distributor district div div1 div2 div3 div4 div5 div6 div7 divGen docAuthor docDate docEdition docImprint docTitle domain eLeaf eTree edition

editionStmt editor editorialDecl education eg egXML elementRef elementSpec ellipsis email emph empty encodingDesc entry entryFree epigraph epilogue equipment equiv etym event eventName ex exemplum expan explicit extent f fDecl fDescr fLib facsimile factuality faith figDesc figure fileDesc filiation finalRubric floatingText floruit foliation foreign forename forest form formula front fs fsConstraints fsDecl fsDescr fsdDecl fsdLink funder fvLib fw g gap gb gen genName gender geo geoDecl geogFeat geogName gi gloss glyph gram gramGrp graph graphic group handDesc handNote handNotes handShift head headItem headLabel height heraldry hi history hom hyph hyphenation iNode iType ident idno if iff imprimatur imprint incident incipit index institution interaction interp interpGrp interpretation item join joinGrp keywords kinesic l label lacunaEnd lacunaStart lang langKnowledge langKnown langUsage language layout layoutDesc lb lbl leaf lem lg licence line link linkGrp list listAnnotation listApp listBibl listChange listEvent listForest listNym listObject listOrg listPerson listPlace listPrefixDef listRef listRelation listTranspose listWit localProp locale location locus locusGrp m macroRef macroSpec mapping material measure measureGrp media meeting memberOf mentioned metDecl metSym metamark milestone mod model modelGrp modelSequence moduleRef moduleSpec monogr mood move msContents msDesc msFrag msIdentifier msItem msItemStruct msName msPart musicNotation name nameLink namespace nationality node normalization notatedMusic note noteGrp notesStmt num number numeric nym oRef object objectDesc objectIdentifier objectName objectType occupation offset opener org orgName orig origDate origPlace origin orth outputRendition p pRef param paramList paramSpec particDesc path pause pb pc per performance persName persPronouns person personGrp persona phr physDesc place placeName population pos post postBox postCode postscript precision prefixDef preparedness principal profileDesc projectDesc prologue pron provenance ptr pubPlace publicationStmt publisher punctuation purpose q quotation quote rb rdg rdgGrp re recordHist recording recordingStmt redo ref refState refsDecl reg region relatedItem relation remarks rendition repository residence resp respStmt respons restore retrace revisionDesc rhyme role roleDesc roleName root row rs rt rubric ruby s said salute samplingDecl schemaRef schemaSpec scriptDesc scriptNote scriptStmt seal sealDesc secFol secl seg segmentation sense sequence series seriesStmt set setting settingDesc settlement sex shift sic signatures signed soCalled socecStatus sound source sourceDesc sourceDoc sp spGrp space span spanGrp speaker specDesc specGrp specGrpRef specList sponsor stage stamp standOff state stdVals street stress string styleDefDecl subc subst substJoin summary superEntry supplied support supportDesc surface surfaceGrp surname surplus surrogates syll symbol table tag tagUsage tagsDecl taxonomy tech teiCorpus teiHeader term terrain text textClass textDesc textLang textNode then time timeline title titlePage titlePart titleStmt tns trailer trait transcriptionDesc transpose tree triangle typeDesc typeNote u unclear undo unicodeProp unihanProp unit unitDecl unitDef usg vAlt vColl vDefault vLabel vMerge vNot vRange val valDesc valItem valList variantEncoding view vocal w watermark when width wit witDetail witEnd witStart witness writing xenoData xr zone]

*Attributes*

`@source` specifies the source from which some aspect of this element is drawn.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Schematron* <sch:rule context="tei:*[@source]"> <sch:let name="srcs" value="tokenize( normalize-space(@source),' ')"/> <sch:report test="( self::tei:classRef | self::tei:dataRef | self::tei:elementRef | self::tei:macroRef | self::tei:moduleRef | self::tei:schemaSpec ) and $srcs[2]"> When used on a schema description element (like <sch:value-of select="name(.)"/>), the @source attribute should have only 1 value. (This one has <sch:value-of select="count($srcs)"/>.) </sch:report> </sch:rule>

*Note* The *source* attribute points to an external source. When used on an element describing a schema component (<classRef>, <dataRef>, <elementRef>, <macroRef>, <moduleRef>, or <schemaSpec>), it identifies the source from which declarations for the components should be obtained.

On other elements it provides a pointer to the bibliographical source from which a quotation or citation is drawn.

In either case, the location may be provided using any form of URI, for example an absolute URI, a relative URI, a private scheme URI of the form `tei:x.y.z`, where

`x.y.z` indicates the version number, e.g. `tei:4.3.2` for TEI P5 release 4.3.2 or (as a special case) `tei:current` for whatever is the latest release, or a private scheme URI that is expanded to an absolute URI as documented in a <prefixDef>.

When used on elements describing schema components, *source* should have only one value; when used on other elements multiple values are permitted.

*Example*

```
<p>
<!-- ... --> As Willard McCarty (<bibl xml:id="mcc_2012">2012, p.2</bibl>)
tells us, <quote source="#mcc_2012">‘Collaboration’ is a problematic and
should be a contested
    term.</quote>
<!-- ... -->
</p>
```

*Example*

```
<p>
<!-- ... -->
  <quote source="#chicago_15_ed">Grammatical theories are in flux, and the
more we learn, the
    less we seem to know.</quote>
<!-- ... -->
</p>
<!-- ... -->
<bibl xml:id="chicago_15_ed">
  <title level="m">The Chicago Manual of Style</title>,
<edition>15th edition</edition>. <pubPlace>Chicago</pubPlace>:
<publisher>University of
    Chicago Press</publisher> (<date>2003</date>),
<biblScope unit="page">p.147</biblScope>.

</bibl>
```

*Example*

```
<elementRef key="p" source="tei:2.0.1"/>
```

Include in the schema an element named <p> available from the TEI P5 2.0.1 release.

*Example*

```
<schemaSpec ident="myODD"
  source="mycompiledODD.xml">
<!-- further declarations specifying the components required -->
</schemaSpec>
```

Create a schema using components taken from the file `mycompiledODD.xml`.

---

## att.handFeatures provides attributes describing aspects of the hand in which a manuscript is written.

[*12.3.2.1. Document Hands*]

*Module* tei — The TEI Infrastructure

*Members* handNote handShift scriptNote typeNote

*Attributes*

- att.scope
  - *@scope*

`@scribe` gives a name or other identifier for the scribe believed to be responsible for this hand.
*Status* Optional
*Datatype* teidata.name

`@scribeRef` points to a full description of the scribe concerned, typically supplied by a <person> element elsewhere in the description.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

`@script` characterizes the particular script or writing style used by this hand, for example *secretary*, *copperplate*, *Chancery*, *Italian*, etc.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.name separated by whitespace

`@scriptRef` points to a full description of the script or writing style used by this hand, typically supplied by a <scriptNote> element elsewhere in the description.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

`@medium` describes the tint or type of ink, e.g. *brown*, or other writing medium, e.g. *pencil*.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace

Usually either *script* or *scriptRef*, and similarly, either *scribe* or *scribeRef*, will be supplied.

---

## att.identified

att.identified provides identifying attributes for elements which can be subsequently referenced by means of a *key* attribute.

*Module* tagdocs — Documentation Elements

*Members* attDef classSpec constraintSpec dataSpec elementSpec macroSpec moduleSpec paramSpec schemaSpec

*Attributes*
- att.combinable
  - *@mode*
- att.deprecated
  - *@validUntil*

`@ident` supplies the identifier by which this element may be referenced.
*Status* Required
*Datatype* teidata.name

`@predeclare` says whether this object should be predeclared in the `tei` infrastructure module.
*Status* Optional
*Datatype* teidata.truthValue
*Default* false

`@module` supplies a name for the module in which this object is to be declared.
*Status* Optional
*Datatype* teidata.xmlName

*Schematron* <sch:rule context="tei:elementSpec[@module]|tei:classSpec[@module]|tei:macroSpec[@module]"> <sch:assert test="(not(ancestor::tei:schemaSpec | ancestor::tei:TEI | ancestor::tei:teiCorpus)) or (not(@module) or (not(//tei:moduleSpec) and not(//tei:moduleRef)) or (//tei:moduleSpec[@ident = current()/@module]) or (//tei:moduleRef[@key = current()/@module]))"> Specification <sch:value-of select="@ident"/>: the value of the module attribute ("<sch:value-of select="@module"/>") should correspond to an existing module, via a moduleSpec or moduleRef</sch:assert> </sch:rule>

---

**att.indentation** provides attributes for describing the indentation of a textual element on the source page or object. [*9.3.2. Attributes Specific to CMC post*]

*Module* tei — The TEI Infrastructure

*Members* post

*Attributes*

`@indentLevel` specifies the level of indentation of an item using a numeric value.

*Status* Recommended

*Datatype* integer

*Note* This attribute can be used to mark the level of indentation of a post in a thread-like structure (as defined by its author and in relation to the standard level of indentation which should be encoded with 0). This is commonly used for Wiki talk pages, but should also be used for genres such as webblog comments when the source was HTML.

---

**att.internetMedia** provides attributes for specifying the type of a computer resource using a standard taxonomy.

*Module* tei — The TEI Infrastructure

*Members* att.media[binaryObject graphic media] equiv ptr ref

*Attributes*

`@mimeType` (MIME media type) specifies the applicable multimedia internet mail extension (MIME) media type.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

*Example* In this example *mimeType* is used to indicate that the URL points to a TEI XML file encoded in UTF-8.

```
<ref  mimeType="application/tei+xml; charset=UTF-8"
  target="https://raw.githubusercontent.com/TEIC/TEI/dev/P5/Source/guidelines-en.xml"/>
```

*Note* This attribute class provides an attribute for describing a computer resource, typically available over the internet, using a value taken from a standard taxonomy. At present only a single taxonomy is supported, the Multipurpose Internet Mail Extensions (MIME) Media Type system. This typology of media types is defined by the Internet Engineering Task Force in RFC 2046. The list of types is maintained by the Internet Assigned Numbers Authority (IANA). The *mimeType* attribute must have a value taken from this list.

---

**att.interpLike** provides attributes for elements which represent a formal analysis or interpretation. [*18.2. Global Attributes for Simple Analyses*]

*Module* tei — The TEI Infrastructure

*Members* interp interpGrp span spanGrp

*Attributes*

`@type` indicates what kind of phenomenon is being noted in the passage.

*Status* Recommended

*Datatype* teidata.enumerated

*Sample values include:* **image** identifies an image in the passage.

**character** identifies a character associated with the passage.

**theme** identifies a theme in the passage.

**allusion** identifies an allusion to another text.

`@subtype` (subtype) provides a sub-categorization of the phenomenon is being noted in the passage, if needed.
*Status* Optional
*Datatype* teidata.enumerated
*Note* The *subtype* attribute may be used to provide any sub-classification for the element additional to that provided by its *type* attribute.

`@inst` (instances) points to instances of the analysis or interpretation represented by the current element.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace
*Note* The current element should be an analytic one. The element pointed at should be a textual one.

---

**att.lexicographic** provides a set of attributes for specifying standard and normalized values, grammatical functions, alternate or equivalent forms, and information about composite parts. [*10.2. The Structure of Dictionary Entries*]

*Module* dictionaries — Dictionaries

*Members* case colloc def entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number oRef orth pRef per pos pron re sense subc syll tns usg xr

*Attributes*
- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*
- att.lexicographic.normalized
  - *@norm*
  - *@orig*

`@expand` (expand) gives an expanded form of information presented more concisely in the dictionary.
*Status* Optional
*Datatype* teidata.text

```
<gramGrp>
  <pos  expand="noun">n</pos>
</gramGrp>
```

`@split` (split) gives the list of split values for a merged form.
*Status* Optional
*Datatype* teidata.text

`@value` (value) gives a value which lacks any realization in the printed source text.
*Status* Optional
*Datatype* teidata.text

`@location` (location) indicates an <anchor> element typically elsewhere in the document, but possibly in another document, which is the original location of this component.
*Status* Optional
*Datatype* teidata.pointer

`@mergedIn` (merged into) gives a reference to another element, where the original appears as a merged form.
*Status* Optional
*Datatype* teidata.pointer

`@opt` (optional) indicates whether the element is optional or not.

*Status* Optional
*Datatype* teidata.truthValue
*Default* false

---

**att.lexicographic.normalized** provides attributes for usage within word-level elements in the analysis module and within lexicographic microstructure in the dictionaries module.

*Module* analysis — Simple Analytic Mechanisms

*Members* att.lexicographic[case colloc def entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number oRef orth pRef per pos pron re sense subc syll tns usg xr] att.linguistic[pc w]

*Attributes*

`@norm` (normalized) provides the normalized/standardized form of information present in the source text in a non-normalized form.

*Status* Optional

*Datatype* teidata.text

Normalization of part-of-speech information within a dictionary entry.
```
<gramGrp>
  <pos norm="noun">n</pos>
</gramGrp>
```
```
Normalization of a source form in a tokenized historical corpus.
<s>
  <w>for</w>
  <w norm="virtue's">vertues</w>
  <w>sake</w>
</s>
<s>
  <w norm="persuasion">perswasion</w>
  <w>of</w>
  <w norm="Unity">Vnitie</w>
</s>
Example of normalization from Aviso. Relation oder Zeitung.
Wolfenbüttel, 1609. In: Deutsches Textarchiv. <s>
  <w norm="freiwillig">freywillig</w>
  <pc norm=","
    join="left">/</pc>
  <w norm="unbedrängt">vnbedra□ngt</w>
  <w norm="und">vnd</w>
  <w norm="unverhindert">vnuerhindert</w>
</s>
<w norm="Teil">Theyll</w>
<w norm="Freude">Frewde</w>
```

`@orig` (original) gives the original string or is the empty string when the element does not appear in the source text.

*Status* Optional

*Datatype* teidata.text

Example from a language documentation project of the Mixtepec-Mixtec language (ISO 639-3: 'mix'). This is a use case where speakers spell something incorrectly but we would like to preserve it for any number of reasons, the use of *orig* is essential and could have uses for both the speaker to see past mistakes, researchers to get insight into how untrained speakers write their language instinctually (in contrast to prescribed convention), etc.: `<w orig="ntsa sia'i">ntsasia'i</w>`

```
Example from the EarlyPrint project. Fragment of text where
obvious errors have been corrected but the original forms remain
recorded: <w lemma="he"
```

```
  pos="pns"
  xml:id="b1afj-003-a-0950">he</w>
<w  lemma="have"
  pos="vvz"
  xml:id="b1afj-003-a-0960">hath</w>
<w  lemma="bring"
  pos="vvn"
  xml:id="b1afj-003-a-0970">brought</w>
<w  lemma="forth"
  pos="av"
  xml:id="b1afj-003-a-0980"
  orig="sorth">forth</w>
An example from the EarlyPrint project showing the use of both
norm and orig. The orig attribute preserves the original version
(sometimes with spelling errors, often with printer
abbreviations), the element content resolves printer
abbreviations but retains the original orthography, and the norm
attribute holds normalized values: <w  lemma="commandment"
  pos="n1"
  norm="commandment"
  xml:id="b9avr-018-a-7720"
  orig="commandemēt">commandement</w>
```

It needs to be stressed that the two attributes in this class are meant for strictly lexicographic and linguistic uses, and not for editorial interventions. For the latter, the mechanism based on <choice>, <orig>, and <reg> needs to be employed.

---

## att.linguistic

att.linguistic provides a set of attributes concerning linguistic features of tokens, for usage within token-level elements, specifically <w> and <pc> in the analysis module. [*18.4.2. Lightweight Linguistic Annotation*]

*Module* analysis — Simple Analytic Mechanisms

*Members* pc w

*Attributes* • att.lexicographic.normalized

- *@norm*
- *@orig*

`@lemma` provides a lemma (base form) for the word, typically uninflected and serving both as an identifier (e.g. in dictionary contexts, as a headword), and as a basis for potential inflections.
*Status* Optional
*Datatype* teidata.text

```
<w  lemma="wife">wives</w>
<w  lemma="Arznei">Artzeneyen</w>
```

`@lemmaRef` provides a pointer to a definition of the lemma for the word, for example in an online lexicon.
*Status* Optional
*Datatype* teidata.pointer

```
<w  type="verb"
  lemma="hit"
  lemmaRef="http://www.example.com/lexicon/hitvb.xml">hitt<m  type="suffix">ing</m>
</w>
```

`@pos` (part of speech) indicates the part of speech assigned to a token (i.e. information on whether it is a noun, adjective, or verb), usually according to some official reference vocabulary (e.g. for German: STTS, for English: CLAWS, for Polish: NKJP, etc.).
*Status* Optional

*Datatype* teidata.text

The German sentence 'Wir fahren in den Urlaub.' tagged with the Stuttgart-Tuebingen-Tagset (STTS).

```
<s>
  <w pos="PPER">Wir</w>
  <w pos="VVFIN">fahren</w>
  <w pos="APPR">in</w>
  <w pos="ART">den</w>
  <w pos="NN">Urlaub</w>
  <w pos="$.">.</w>
</s>
```

```
The English sentence 'We're going to Brazil.' tagged with the
CLAWS-5 tagset, arranged inline (with significant whitespace).
<p><w pos="PNP">We</w><w pos="VBB">'re</w> <w pos="VVG">going</w> <w pos="PRP">to<

The English sentence 'We're going on vacation to Brazil for a
month!' tagged with the CLAWS-7 tagset and arranged
sequentially. <p>
  <w pos="PPIS2">We</w>
  <w pos="VBR">'re</w>
  <w pos="VVG">going</w>
  <w pos="II">on</w>
  <w pos="NN1">vacation</w>
  <w pos="II">to</w>
  <w pos="NP1">Brazil</w>
  <w pos="IF">for</w>
  <w pos="AT1">a</w>
  <w pos="NNT1">month</w>
  <pc pos="!">!</pc>
</p>
```

`@msd` (morphosyntactic description) supplies morphosyntactic information for a token, usually according to some official reference vocabulary (e.g. for German: STTS-large tagset; for a feature description system designed as (pragmatically) universal, see Universal Features).

*Status* Optional

*Datatype* teidata.text

```
<ab>
  <w pos="PPER"
    msd="1.Pl.*.Nom">Wir</w>
  <w pos="VVFIN"
    msd="1.Pl.Pres.Ind">fahren</w>
  <w pos="APPR"
    msd="--">in</w>
  <w pos="ART"
    msd="Def.Masc.Akk.Sg">den</w>
  <w pos="NN"
    msd="Masc.Akk.Sg">Urlaub</w>
  <pc pos="$."
    msd="--">.</pc>
</ab>
```

`@join` when present, provides information on whether the token in question is adjacent to another, and if so, on which side.

*Status* Optional

*Datatype* teidata.text

*Legal values are:* **no** the token is not adjacent to another

**left** there is no whitespace on the left side of the token

**right** there is no whitespace on the right side of the token

**both** there is no whitespace on either side of the token

**overlap** the token overlaps with another; other devices (specifying the extent and the area of overlap) are needed to more precisely locate this token in the character stream

The example below assumes that the lack of whitespace is marked redundantly, by using the appropriate values of *join*.

```
<s>
  <pc  join="right">"</pc>
  <w  join="left">Friends</w>
  <w>will</w>
  <w>be</w>
  <w  join="right">friends</w>
  <pc  join="both">.</pc>
  <pc  join="left">"</pc>
</s>Note that a project may make a decision to only indicate
lack of whitespace in one direction, or do that non-redundantly.
The existing proposal is the broadest possible, on the
assumption that we adopt the "streamable view", where all the
information on the current element needs to be represented
locally.
The English sentence ‘We're going on vacation.’ tagged with the
CLAWS-5 tagset, arranged sequentially, tagged on the assumption
that only the lack of the preceding whitespace is indicated. <p>
  <w  pos="PNP">We</w>
  <w  pos="VBB"
    join="left">'re</w>
  <w  pos="VVG">going</w>
  <w  pos="PRP">on</w>
  <w  pos="NN1">vacation</w>
  <pc  pos="PUN"
    join="left">.</pc>
</p>
```

*Note* The definition of this attribute is adapted from ISO MAF (Morpho-syntactic Annotation Framework), ISO 24611:2012.

These attributes make it possible to encode simple language corpora and to add a layer of linguistic information to any tokenized resource. See section *18.4.2. Lightweight Linguistic Annotation* for discussion.

---

**att.locatable** provides attributes for referencing locations by pointing to entries in a canonical list of places. [*2.3.9. The Unit Declaration 14.3.4.3. States, Traits, and Events*]

*Module* tei — The TEI Infrastructure

*Members* conversion event

*Attributes*

`@where` indicates one or more locations by pointing to a <place> element or other canonical description.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

---

**att.measurement** provides attributes to represent a regularized or normalized measurement.

*Module* tei — The TEI Infrastructure

*Members* measure measureGrp unit

*Attributes*

`@unit` (unit) indicates the units used for the measurement, usually using the standard symbol for the desired units.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **m** (metre) SI base unit of length
**kg** (kilogram) SI base unit of mass
**s** (second) SI base unit of time
**Hz** (hertz) SI unit of frequency
**Pa** (pascal) SI unit of pressure or stress
**Ω** (ohm) SI unit of electric resistance
**L** (litre) 1 $dm^3$
**t** (tonne) $10^3$ kg
**ha** (hectare) 1 $hm^2$
**Å** (ångström) $10^{-10}$ m
**mL** (millilitre)
**cm** (centimetre)
**dB** (decibel) see remarks, below
**kbit** (kilobit) $10^3$ or 1000 bits
**Kibit** (kibibit) $2^{10}$ or 1024 bits
**kB** (kilobyte) $10^3$ or 1000 bytes
**KiB** (kibibyte) $2^{10}$ or 1024 bytes
**MB** (megabyte) $10^6$ or 1 000 000 bytes
**MiB** (mebibyte) $2^{20}$ or 1 048 576 bytes
*Note* If the measurement being represented is not expressed in a particular unit, but rather is a number of discrete items, the unit count should be used, or the *unit* attribute may be left unspecified.
Wherever appropriate, a recognized SI unit name should be used (see further `http://www.bipm.org/en/publications/si-brochure/`; `http://physics.nist.gov/cuu/Units/`). The list above is indicative rather than exhaustive.

`@unitRef` points to a unique identifier stored in the *xml:id* of a <unitDef> element that defines a unit of measure.
*Status* Optional
*Datatype* teidata.pointer

`@quantity` (quantity) specifies the number of the specified units that comprise the measurement
*Status* Optional
*Datatype* teidata.numeric

`@commodity` (commodity) indicates the substance that is being measured
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.word separated by whitespace
*Note* In general, when the commodity is made of discrete entities, the plural form should be used, even when the measurement is of only one of them.

*Schematron* <sch:rule context="tei:*[@unitRef]"> <sch:report test="@unit" role="info">The @unit attribute may be unnecessary when @unitRef is present.</sch:report> </sch:rule>

*Note* This attribute class provides a triplet of attributes that may be used either to regularize the values of the measurement being encoded, or to normalize them with respect to a standard measurement system.

```
<l>
<!-- regularization:-->
So weren't you gonna buy <measure  quantity="0.5"  unit="gal"
   commodity="ice cream">half
    a gallon</measure>, baby
</l>
<l>
<!-- normalization: -->
So won't you go and buy <measure  quantity="1.893"  unit="L"
   commodity="ice cream">half
    a gallon</measure>, baby?
</l>
```

The unit should normally be named using the standard symbol for an SI unit (see further `http://www.bipm.org/en/publications/si-brochure/`; `http://physics.nist.gov/cuu/Units/`). However, encoders may also specify measurements using informally defined units such as lines or characters.

---

**att.media** provides attributes for specifying display and related properties of external media.

*Module* tei — The TEI Infrastructure

*Members* binaryObject graphic media

*Attributes* • att.internetMedia
- *@mimeType*

`@width` Where the media are displayed, indicates the display width.
*Status* Optional
*Datatype* teidata.outputMeasurement

`@height` Where the media are displayed, indicates the display height.
*Status* Optional
*Datatype* teidata.outputMeasurement

`@scale` Where the media are displayed, indicates a scale factor to be applied when generating the desired display size.
*Status* Optional
*Datatype* teidata.numeric

---

**att.metrical** defines a set of attributes that certain elements may use to represent metrical information. [*6.4. Rhyme and Metrical Analysis*]

*Module* verse — Verse

*Members* att.divLike[div div1 div2 div3 div4 div5 div6 div7 lg] att.segLike[c cl m pc phr s seg w] l

*Attributes*

`@met` (metrical structure, conventional) contains a user-specified encoding for the conventional metrical structure of the element.
*Status* Recommended
*Datatype* token
*Note* The pattern may be specified by means of either a standard term for the kind of metrical unit (e.g. *hexameter*) or an encoded representation for the metrical pattern (e.g. *+−+-+-+-+-*). In either case, the notation used should be documented by a <metDecl> element within the <encodingDesc> of the associated header.

Where this attribute is not specified, the metrical pattern for the element concerned is understood to be inherited from its parent.

`@real` (metrical structure, realized) contains a user-specified encoding for the actual realization of the conventional metrical structure applicable to the element.

*Status* Optional

*Datatype* token

*Note* The pattern may be specified by means of either a standard term for the kind of metrical unit (e.g. *hexameter*) or an encoded representation for the metrical pattern (e.g. *+−+-+-+-+-*). In either case, the notation used should be documented by a <metDecl> element within the <encodingDesc> of the associated header.

Where this attribute is not specified, the metrical realization for the element concerned is understood to be identical to that specified or implied for the *met* attribute.

`@rhyme` (rhyme scheme) specifies the rhyme scheme applicable to a group of verse lines.

*Status* Recommended

*Datatype* token

*Note* By default, the rhyme scheme is expressed as a string of alphabetic characters each corresponding with a rhyming line. Any non-rhyming lines should be represented by a hyphen or an X. Alternative notations may be defined as for *met* by use of the <metDecl> element in the TEI header.

When the default notation is used, it does not make sense to specify this attribute on any unit smaller than a line. Nor does the default notation provide any way to record internal rhyme, or to specify non-conventional rhyming practice. These extensions would require user-defined alternative notations.

---

**att.milestoneUnit** provides attributes to indicate the type of section which is changing at a specific milestone. [*3.11.3. Milestone Elements 2.3.6.3. Milestone Method 2.3.6. The Reference System Declaration*]

*Module* core — Elements Available in All TEI Documents

*Members* milestone refState

*Attributes*

`@unit` provides a conventional name for the kind of section changing at this milestone.

*Status* Required

*Datatype* teidata.enumerated

*Suggested values include:* **page** physical page beginnings (synonymous with the <pb> element).

**column** column beginnings.

**line** line beginnings (synonymous with the <lb> element).

**book** any units termed book, liber, etc.

**poem** individual poems in a collection.

**canto** cantos or other major sections of a poem.

**speaker** changes of speaker or narrator.

**stanza** stanzas within a poem, book, or canto.

**act** acts within a play.

**scene** scenes within a play or act.

**section** sections of any kind.

**absent** passages not present in the reference edition.

**unnumbered** passages present in the text, but not to be included as part of the reference.

```
<milestone n="23"
  ed="La"
  unit="Dreissiger"/>
... <milestone n="24"
  ed="AV"
  unit="verse"/> ...
```

*Note* If the milestone marks the beginning of a piece of text not present in the reference edition, the special value *absent* may be used as the value of *unit*. The normal interpretation is that the reference edition does not contain the text which follows, until the next <milestone> tag for the edition in question is encountered.

In addition to the values suggested, other terms may be appropriate (e.g. *Stephanus* for the Stephanus numbers in Plato).

The *type* attribute may be used to characterize the unit boundary in any respect other than simply identifying the type of unit, for example as word-breaking or not.

---

**att.msClass** provides attributes to indicate text type or classification. [*11.6. Intellectual Content 11.6.1. The msItem and msItemStruct Elements*]

*Module* msdescription — Manuscript Description

*Members* msContents msItem msItemStruct

*Attributes*

`@class` identifies the text types or classifications applicable to this item by pointing to other elements or resources defining the classification concerned.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

---

**att.msExcerpt** (manuscript excerpt) provides attributes used to describe excerpts from a manuscript placed in a description thereof. [*11.6. Intellectual Content*]

*Module* msdescription — Manuscript Description

*Members* colophon explicit finalRubric incipit msContents msItem msItemStruct quote rubric

*Attributes*

`@defective` indicates whether the passage being quoted is defective, i.e. incomplete through loss or damage.

*Status* Optional

*Datatype* teidata.xTruthValue

In the case of an incipit, indicates whether the incipit as given is defective, i.e. the first words of the text as preserved, as opposed to the first words of the work itself. In the case of an explicit, indicates whether the explicit as given is defective, i.e. the final words of the text as preserved, as opposed to what the closing words would have been had the text of the work been whole.

---

**att.namespaceable** provides attributes indicating the target namespace for an object being created.

*Module* tagdocs — Documentation Elements

*Members* attDef elementSpec schemaSpec

*Attributes*

`@ns` (namespace) specifies the namespace to which the element(s) being specified belongs.

*Status* Optional

*Datatype* 0–1 occurrences of teidata.namespace separated by whitespace

*Default* http://www.tei-c.org/ns/1.0

*Note* When specified on an <elementSpec>, this attribute specifies the namespace of the element being added, deleted, replaced, or changed. When specified on a <schemaSpec>, this attribute specifies the namespace of every element that is added, deleted, replaced, or changed by <elementSpec> children of the <schemaSpec>, or by <elementSpec>s referred to by <specGrpRef> children of the <schemaSpec>, *unless* overridden by an *ns* attribute on the <elementSpec>.

Note that the value of *ns* of <attDef> is *not* inherited from its ancestors.

---

**att.naming** provides attributes common to elements which refer to named persons, places, organizations etc. [*3.6.1. Referring Strings 14.3.7. Names and Nyms*]

*Module* tei — The TEI Infrastructure

*Members* att.personal[addName eventName forename genName name objectName orgName persName placeName roleName surname] affiliation author birth bloc climate collection country death district editor education event geogFeat geogName institution nationality occupation offset origPlace population pubPlace region repository residence rs settlement socecStatus state terrain trait

*Attributes*

- att.canonical
  - *@key*
  - *@ref*

`@role` may be used to specify further information about the entity referenced by this name in the form of a set of whitespace-separated values, for example the occupation of a person, or the status of a place.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace

`@nymRef` (reference to the canonical name) provides a means of locating the canonical form (*nym*) of the names associated with the object named by the element bearing it.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Note* The value must point directly to one or more XML elements by means of one or more URIs, separated by whitespace. If more than one is supplied, the implication is that the name is associated with several distinct canonical names.

---

**att.notated** provides attributes to indicate any specialised notation used for element content.

*Module* tei — The TEI Infrastructure

*Members* c cl formula hyph listAnnotation m oRef orth pRef phr pron quote s seg stress syll u w

*Attributes*

`@notation` names the notation used for the content of the element.

*Status* Optional

*Datatype* teidata.enumerated

---

**att.partials** provides attributes for describing the extent of lexical references for a dictionary term.

*Module* tei — The TEI Infrastructure

*Members* orth pron

*Attributes*

`@extent` indicates whether the pronunciation or orthography applies to all or part of a word.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **full** (full form)
**pref** (prefix)
**suff** (suffix)
**inf** (infix)
**part** (partial)
*Note* This attribute is optional, and no default value is specified, so it can be omitted if this information is not necessary.

---

**att.patternReplacement** provides attributes for regular-expression matching and replacement. [*17.2.3. Using Abbreviated Pointers 2.3.6.3. Milestone Method 2.3.6. The Reference System Declaration 2.3.6.2. Search-and-Replace Method*]

*Module* header — The TEI Header

*Members* cRefPattern prefixDef

*Attributes*

`@matchPattern` specifies a regular expression against which the values of other attributes can be matched.
*Status* Required
*Datatype* teidata.pattern
*Note* The syntax used should follow that defined by W3C XPath syntax. Note that parenthesized groups are used not only for establishing order of precedence and atoms for quantification, but also for creating subpatterns to be referenced by the *replacementPattern* attribute.

`@replacementPattern` specifies a 'replacement pattern', that is, the skeleton of a relative or absolute URI containing references to groups in the *matchPattern* which, once subpattern substitution has been performed, complete the URI.
*Status* Required
*Datatype* teidata.replacement
*Note* The strings *$1*, *$2* etc. are references to the corresponding group in the regular expression specified by *matchPattern* (counting open parenthesis, left to right). Processors are expected to replace them with whatever matched the corresponding group in the regular expression.
If a digit preceded by a dollar sign is needed in the actual replacement pattern (as opposed to being used as a back reference), the dollar sign must be written as `%24`.

---

**att.personal** (attributes for components of names usually, but not necessarily, personal names) common attributes for those elements which form part of a name usually, but not necessarily, a personal name. [*14.2.1. Personal Names*]

*Module* tei — The TEI Infrastructure

*Members* addName eventName forename genName name objectName orgName persName placeName roleName surname

*Attributes*
- att.naming
  - *@role*

– *@nymRef*

– att.canonical
* *@key*
* *@ref*

`@full` indicates whether the name component is given in full, as an abbreviation or simply as an initial.
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **yes** (yes) the name component is spelled out in full.[Default]
**abb** (abbreviated) the name component is given in an abbreviated form.
**init** (initial letter) the name component is indicated only by one initial.

`@sort` (sort) specifies the sort order of the name component in relation to others within the name.
*Status* Optional
*Datatype* teidata.count

---

**att.placement** provides attributes for describing where on the source page or object a textual element appears. [*3.5.3. Additions, Deletions, and Omissions 12.3.1.4. Additions and Deletions*]

*Module* tei — The TEI Infrastructure

*Members* att.transcriptional[add addSpan del delSpan mod redo restore retrace rt subst substJoin undo] div figure fw head label metamark notatedMusic note noteGrp postscript sp speaker stage trailer witDetail

*Attributes*

`@place` specifies where this item is placed.
*Status* Recommended
*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace
*Suggested values include:* **top** at the top of the page
**bottom** at the foot of the page
**margin** in the margin (left, right, or both)
**opposite** on the opposite, i.e. facing, page
**overleaf** on the other side of the leaf
**above** above the line
**right** to the right, e.g. to the right of a vertical line of text, or to the right of a figure
**below** below the line
**left** to the left, e.g. to the left of a vertical line of text, or to the left of a figure
**end** at the end of e.g. chapter or volume.
**inline** within the body of the text.
**inspace** in a predefined space, for example left by an earlier scribe.

```
<add place="margin">[An addition written in the margin]</add>
<add place="bottom opposite">[An addition written at the
foot of the current page and also on the facing page]</add>
<note place="bottom">Ibid, p.7</note>
```

---

**att.pointing** provides a set of attributes used by all elements which point to other elements by means of one or more URI references. [*1.3.1.1.2. Language Indicators 3.7. Simple Links and Cross-References*]

*Module* tei — The TEI Infrastructure

*Members* att.pointing.group[altGrp joinGrp linkGrp] alt annotation calendar catRef citedRange gloss join licence link locus note noteGrp oRef pRef ptr ref span substJoin term witDetail

*Attributes*

`@targetLang` specifies the language of the content to be found at the destination referenced by *target*, using a 'language tag' generated according to BCP 47.

*Status* Optional

*Datatype* teidata.language

*Schematron* <sch:rule context="tei:*[not(self::tei:schemaSpec)][@targetLang]"> <sch:assert test="@target">@targetLang should only be used on <sch:name/> if @target is specified.</sch:assert> </sch:rule>

```
<linkGrp xml:id="pol-swh_aln_2.1-linkGrp">
  <ptr xml:id="pol-swh_aln_2.1.1-ptr"
    target="pol/UDHR/text.xml#pol_txt_1-head"
    type="tuv"
    targetLang="pl"/>
  <ptr xml:id="pol-swh_aln_2.1.2-ptr"
    target="swh/UDHR/text.xml#swh_txt_1-head"
    type="tuv"
    targetLang="sw"/>
</linkGrp>In the example above, the <linkGrp> combines pointers
at parallel fragments of the Universal Declaration of Human
Rights: one of them is in Polish, the other in Swahili.
```

*Note* The value must conform to BCP 47. If the value is a private use code (i.e., starts with x- or contains -x-), a <language> element with a matching value for its *ident* attribute should be supplied in the TEI header to document this value. Such documentation may also optionally be supplied for non-private-use codes, though these must remain consistent with their ( IETF) Internet Engineering Task Force definitions.

`@target` specifies the destination of the reference by supplying one or more URI References.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Note* One or more syntactically valid URI references, separated by whitespace. Because whitespace is used to separate URIs, no whitespace is permitted inside a single URI. If a whitespace character is required in a URI, it should be escaped with the normal mechanism, e.g. `TEI%20Consortium`.

`@evaluate` (evaluate) specifies the intended meaning when the target of a pointer is itself a pointer.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **all** if the element pointed to is itself a pointer, then the target of that pointer will be taken, and so on, until an element is found which is not a pointer.

**one** if the element pointed to is itself a pointer, then its target (whether a pointer or not) is taken as the target of this pointer.

**none** no further evaluation of targets is carried out beyond that needed to find the element specified in the pointer's target.

*Note* If no value is given, the application program is responsible for deciding (possibly on the basis of user input) how far to trace a chain of pointers.

---

**att.pointing.group** provides a set of attributes common to all elements which enclose groups of pointer elements. [*17. Linking, Segmentation, and Alignment*]

*Module* tei — The TEI Infrastructure
*Members* altGrp joinGrp linkGrp
*Attributes*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*
- att.typed
  - *@type*
  - *@subtype*

`@domains` optionally specifies the identifiers of the elements within which all elements indicated by the contents of this element lie.
*Status* Optional
*Datatype* 2–∞ occurrences of teidata.pointer separated by whitespace
*Note* If this attribute is supplied every element specified as a target must be contained within the element or elements named by it. An application may choose whether or not to report failures to satisfy this constraint as errors, but may not access an element of the right identifier but in the wrong context. If this attribute is not supplied, then target elements may appear anywhere within the target document.

`@targFunc` (target function) describes the function of each of the values of the *target* attribute of the enclosed <link>, <join>, or <alt> tags.
*Status* Optional
*Datatype* 2–∞ occurrences of teidata.word separated by whitespace
*Note* The number of separate values must match the number of values in the *target* attribute in the enclosed <link>, <join>, or <alt> tags (an intermediate <ptr> element may be needed to accomplish this). It should also match the number of values in the *domains* attribute, of the current element, if one has been specified.

---

## att.predicate provides attributes for filtering by an XPath predicate expression. [*23.4. Common Elements 23.5.4.1. The TEI processing model 23.5.4.9. Implementation of Processing Models*]

*Module* tagdocs — Documentation Elements
*Members* equiv model
*Attributes*

`@predicate` the condition under which the element bearing this attribute applies, given as an XPath predicate expression.
*Status* Optional
*Datatype* teidata.xpath
*Note* The XPath predicate expression given as the value of the *predicate* attribute has to be provided *without* wrapping square brackets.

*Example*

```
<model predicate="parent::person"
  behaviour="inline">
  <desc versionDate="2015-08-21"
    xml:lang="en">If it is a child of a person element, treat as inline</desc>
</model>
```

*Example* The following example declares that the <name> element can be mapped to, or is equivalent to, the external concepts of 'PERSON' and 'PLACE' depending on the 'XPath' expression given in *predicate*

```
<elementSpec ident="name" mode="change">
  <equiv name="PERSON"
    predicate="@type eq 'person'"
    uri="http://www.example.com/entities/person"/>
  <equiv name="PLACE"
    predicate="@type eq 'place'"
    uri="http://www.example.com/entities/place"/>
</elementSpec>
```

---

**att.ranging** provides attributes for describing numerical ranges.

*Module* tei — The TEI Infrastructure

*Members* att.dimensions[att.damaged[damage damageSpan] add addSpan age birth date death del delSpan depth dim dimensions ellipsis ex floruit gap geogFeat height mod offset origDate population redo restore retrace secl space state subst substJoin supplied surplus time trait unclear undo width] measure num precision

*Attributes*

`@atLeast` gives a minimum estimated value for the approximate measurement.
*Status* Optional
*Datatype* teidata.numeric

`@atMost` gives a maximum estimated value for the approximate measurement.
*Status* Optional
*Datatype* teidata.numeric

`@min` where the measurement summarizes more than one observation or a range, supplies the minimum value observed.
*Status* Optional
*Datatype* teidata.numeric

`@max` where the measurement summarizes more than one observation or a range, supplies the maximum value observed.
*Status* Optional
*Datatype* teidata.numeric

`@confidence` specifies the degree of statistical confidence (between zero and one) that a value falls within the range specified by *min* and *max*, or the proportion of observed values that fall within that range.
*Status* Optional
*Datatype* teidata.probability

*Example*

```
 The MS. was lost in transmission by mail from
<del rend="overstrike">
  <gap reason="illegible"
    extent="one or two letters" atLeast="1" atMost="2" unit="chars"/>
</del>
Philadelphia to the Graphic office, New York.
```

*Example*

```
 Americares has been supporting the health sector in Eastern Europe since 1986,
and since 1992 has provided <measure atLeast="120000000" unit="USD"
```

```
      commodity="currency">more
    than $120m</measure> in aid to Ukrainians.
```

---

**att.rdgPart** provides attributes to mark the beginning or ending of a fragmentary manuscript or other witness. [*13.1.5. Fragmentary Witnesses*]

*Module* textcrit — Critical Apparatus

*Members* lacunaEnd lacunaStart wit witEnd witStart

*Attributes*

`@wit` (witness or witnesses) contains a space-delimited list of one or more sigla indicating the witnesses to this reading beginning or ending at this point.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

These elements may appear anywhere within the elements <lem> and <rdg>, and also within any of their constituent elements.

---

**att.repeatable** provides attributes for the elements which define component parts of a content model.

*Module* tagdocs — Documentation Elements

*Members* alternate anyElement classRef datatype elementRef sequence

*Attributes*

`@minOccurs` (minimum number of occurrences) indicates the smallest number of times this component may occur.
*Status* Optional
*Datatype* teidata.count
*Default* 1

`@maxOccurs` (maximum number of occurrences) indicates the largest number of times this component may occur.
*Status* Optional
*Datatype* teidata.unboundedCount
*Default* 1

*Schematron* <sch:rule context="tei:*[ @minOccurs and @maxOccurs ]"> <sch:let name="min" value="@minOccurs cast as xs:integer"/> <sch:let name="max" value="if ( normalize-space( @maxOccurs ) eq 'unbounded') then -1 else @maxOccurs cast as xs:integer"/> <sch:assert test="$max eq -1 or $max ge $min">@maxOccurs should be greater than or equal to @minOccurs</sch:assert> </sch:rule> <sch:rule context="tei:*[ @minOccurs and not( @maxOccurs ) ]"> <sch:assert test="@minOccurs cast as xs:integer lt 2">When @maxOccurs is not specified, @minOccurs must be 0 or 1</sch:assert> </sch:rule>

*Note* The value of *minOccurs* must always be less than or equal to that of *maxOccurs*. Since the default value of *maxOccurs* is 1, when *maxOccurs* is not specified *minOccurs* must always be less than or equal to 1. The default value of *minOccurs* is also 1, and therefore, when *minOccurs* is not specified the value of *maxOccurs* must always be greater than or equal to 1. An ODD processor should raise an error if either of these conditions is not met.

---

**att.resourced** provides attributes by which a resource (such as an externally held media file) may be located.

*Module* tei — The TEI Infrastructure

*Members* graphic media schemaRef

*Attributes*

`@url` (uniform resource locator) specifies the URL from which the media concerned may be obtained.

*Status* Required

*Datatype* teidata.pointer

---

**att.scope** provides attributes to describe, in general terms, the scope of an element's application.

*Module* tei — The TEI Infrastructure

*Members* att.handFeatures[handNote handShift scriptNote typeNote] language

*Attributes*

`@scope` indicates the scope of application of the element

*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **sole** only this particular feature is used throughout the document

**major** this feature is used through most of the document

**minor** this feature is used occasionally through the document

```
<langUsage>
  <language ident="en"
    scope="major"/>
  <language ident="es"
    scope="minor"/>
  <language ident="x-ww"
    scope="minor">An invented language the children call
<name>Wikwah</name>.</language>
</langUsage>
<handNote scope="sole">
  <p>Written in insular phase II half-uncial with
     interlinear Old English gloss in an Anglo-Saxon
     pointed minuscule.</p>
</handNote>
```

---

**att.scoping** provides attributes for selecting particular elements within a document.

*Module* tei — The TEI Infrastructure

*Members* certainty precision respons

*Attributes*

`@target` points at one or more sets of zero or more elements each.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

```
Elizabeth went to <persName xml:id="ESSEX">Essex</persName>
<certainty target="#ESSEX"
  locus="name"
  degree="0.6"/>
```

`@match` supplies an XPath selection pattern using the syntax defined in **?** which identifies a set of nodes, selected within the context identified by the *target* attribute if this is supplied, or within the context of the parent element if it is not.

*Status* Optional

*Datatype* teidata.xpath

```
<gap  reason="cancelled">
  <certainty  match="@reason"
    locus="value"
    cert="low"/>
</gap>
```

The semantics of this element apply to the nodeset identified by the value of the *target* attribute, possibly modified by the value of the *match* attribute. If more than one identifier is given, the implication is that all elements (or nodesets) are intended. The *match* attribute may also be used as a means of identifying groups of elements.

If *target* and *match* are present, *target* selects an element and the XPath expression in *match* is evaluated in the context of that element. If neither attribute is present, the expression applies to its parent element. If only *target* is given, the expression refers to the selected element or nodeset. If only *match* is given, the XPath expression is evaluated in the context of the parent element of the bearing element.

Note that the value of the *target* attribute may include an XPointer expression including an XPath expression (see *17.2.4. TEI XPointer Schemes*).

---

**att.segLike** provides attributes for elements used for arbitrary segmentation. [*17.3. Blocks, Segments, and Anchors 18.1. Linguistic Segment Categories*]

*Module* tei — The TEI Infrastructure

*Members* c cl m pc phr s seg w

*Attributes*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*
- att.fragmentable
  - *@part*
- att.metrical
  - *@met*
  - *@real*
  - *@rhyme*

`@function` (function) characterizes the function of the segment.

*Status* Optional

*Datatype* teidata.enumerated

*Note* Attribute values will often vary depending on the type of element to which they are attached. For example, a <cl>, may take values such as coordinate, subject, adverbial etc. For a <phr>, such values as subject, predicate etc. may be more appropriate. Such constraints will typically be implemented by a project-defined customization.

---

**att.sortable** provides attributes for elements in lists or groups that are sortable, but whose sorting key cannot be derived mechanically from the element content. [*10.1. Dictionary Body and Overall Structure*]

*Module* tei — The TEI Infrastructure

*Members* bibl biblFull biblStruct correspAction entry entryFree event idno item list listApp listBibl listChange listEvent listNym listObject listOrg listPerson listPlace listRelation listWit msDesc nym object org person personGrp persona place relation superEntry term witness

*Attributes*

`@sortKey` supplies the sort key for this element in an index, list or group which contains it.

*Status* Optional

*Datatype* teidata.word

```
David's other principal backer, Josiah
ha-Kohen <index indexName="NAMES">
  <term sortKey="Azarya_Josiah_Kohen">Josiah ha-Kohen b.
Azarya</term>
</index> b. Azarya, son of one of the last gaons of Sura was
David's own first
cousin.
```

*Note* The sort key is used to determine the sequence and grouping of entries in an index. It provides a sequence of characters which, when sorted with the other values, will produced the desired order; specifics of sort key construction are application-dependent

Dictionary order often differs from the collation sequence of machine-readable character sets; in English-language dictionaries, an entry for *4-H* will often appear alphabetized under 'fourh', and *McCoy* may be alphabetized under 'maccoy', while *A1*, *A4*, and *A5* may all appear in numeric order 'alphabetized' between 'a-' and 'AA'. The sort key is required if the orthography of the dictionary entry does not suffice to determine its location.

---

## att.spanning provides attributes for elements which delimit a span of text by pointing mechanisms rather than by enclosing it. [*12.3.1.4. Additions and Deletions 1.3.1. Attribute Classes*]

*Module* tei — The TEI Infrastructure

*Members* addSpan cb damageSpan delSpan gb index lb metamark milestone mod pb redo retrace undo

*Attributes*

`@spanTo` indicates the end of a span initiated by the element bearing this attribute.

*Status* Optional

*Datatype* teidata.pointer

*Schematron* The @spanTo attribute must point to an element following the current element; however, this can only be tested if both this element and the one pointed to are in the same document. <sch:rule context="tei:*[ starts-with( @spanTo, '#') ]"> <sch:assert test="id( substring( @spanTo, 2 ) ) » .">The element indicated by @spanTo (<sch:value-of select="@spanTo"/>) must follow the current <sch:name/> element </sch:assert> </sch:rule>

The span is defined as running in document order from the start of the content of the pointing element to the end of the content of the element pointed to by the *spanTo* attribute (if any). If no value is supplied for the attribute, the assumption is that the span is coextensive with the pointing element. If no content is present, the assumption is that the starting point of the span is immediately following the element itself.

---

## att.styleDef provides attributes to specify the name of a formal definition language used to provide formatting or rendition information.

*Module* tei — The TEI Infrastructure

*Members* rendition styleDefDecl

*Attributes*

`@scheme` identifies the language used to describe the rendition.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **css** Cascading Stylesheet Language

**xslfo** Extensible Stylesheet Language Formatting Objects

**free** Informal free text description

**other** A user-defined rendition description language

*Note* If no value for the @scheme attribute is provided, then the default assumption should be that CSS is in use.

`@schemeVersion` supplies a version number for the style language provided in *scheme*.

*Status* Optional

*Datatype* teidata.versionNumber

*Schematron* <sch:rule context="tei:*[@schemeVersion]"> <sch:assert test="@scheme and not(@scheme = 'free')"> @schemeVersion can only be used if @scheme is specified. </sch:assert> </sch:rule>

*Note* If *schemeVersion* is used, then *scheme* should also appear, with a value other than free.

---

**att.tableDecoration** provides attributes used to decorate rows or cells of a table. [*15. Tables, Formulæ, Graphics, and Notated Music*]

*Module* figures — Tables, Formulæ, Graphics, and Notated Music

*Members* cell row

*Attributes*

`@role` (role) indicates the kind of information held in this cell or in each cell of this row.

*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **label** labelling or descriptive information only.

**data** data values.[Default]

*Note* When this attribute is specified on a row, its value is the default for all cells in this row. When specified on a cell, its value overrides any default specified by the *role* attribute of the parent <row> element.

`@rows` (rows) indicates the number of rows occupied by this cell or row.

*Status* Optional

*Datatype* teidata.count

*Default* 1

*Note* A value greater than one indicates that this cell spans several rows. Where several cells span multiple rows, it may be more convenient to use nested tables.

`@cols` (columns) indicates the number of columns occupied by this cell or row.

*Status* Optional

*Datatype* teidata.count

*Default* 1

*Note* A value greater than one indicates that this cell or row spans several columns. Where an initial cell spans an entire row, it may be better treated as a heading.

---

**att.textCritical** defines a set of attributes common to all elements representing variant readings in text critical work. [*13.1. The Apparatus Entry, Readings, and Witnesses*]

*Module* textcrit — Critical Apparatus

*Members* lem rdg rdgGrp

*Attributes* • att.written
- *@hand*

• att.typed
- ~~type~~
- @subtype

`@type` classifies the reading according to some useful typology.
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **substantive** (substantive) the reading offers a substantive variant.
**orthographic** (orthographic) the reading differs only orthographically, not in substance, from other readings.

`@cause` classifies the cause for the variant reading, according to any appropriate typology of possible origins.
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **homeoteleuton**
**homeoarchy**
**paleographicConfusion**
**haplography**
**dittography**
**falseEmendation**

`@varSeq` (variant sequence) provides a number indicating the position of this reading in a sequence, when there is reason to presume a sequence to the variants.
*Status* Optional
*Datatype* teidata.count
*Note* Different variant sequences could be coded with distinct number trails: 1-2-3 for one sequence, 5-6-7 for another. More complex variant sequences, with (for example) multiple branchings from single readings, may be expressed through the <join> element.

`@require` points to other readings that are required when adopting the current reading or lemma.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

This element class defines attributes inherited by <rdg>, <lem>, and <rdgGrp>.

---

## att.timed

**att.timed** provides attributes common to those elements which have a duration in time, expressed either absolutely or by reference to an alignment map. [*8.3.5. Temporal Information*]

*Module* tei — The TEI Infrastructure

*Members* annotationBlock binaryObject ellipsis gap incident kinesic media pause post u vocal writing

*Attributes* • att.duration
- att.duration.iso
  * *@dur-iso*
- att.duration.w3c
  * *@dur*

`@start` indicates the location within a temporal alignment at which this element begins.
*Status* Optional
*Datatype* teidata.pointer

*Note* If no value is supplied, the element is assumed to follow the immediately preceding element at the same hierarchic level.

`@end` indicates the location within a temporal alignment at which this element ends.

*Status* Optional

*Datatype* teidata.pointer

*Note* If no value is supplied, the element is assumed to precede the immediately following element at the same hierarchic level.

---

**att.transcriptional** provides attributes specific to elements encoding authorial or scribal intervention in a text when transcribing manuscript or similar sources. [*12.3.1.4. Additions and Deletions*]

*Module* tei — The TEI Infrastructure

*Members* add addSpan del delSpan mod redo restore retrace rt subst substJoin undo

*Attributes*

- att.editLike
  - *@evidence*
  - *@instant*
- att.placement
  - *@place*
- att.written
  - *@hand*

`@status` indicates the effect of the intervention, for example in the case of a deletion, strikeouts which include too much or too little text, or in the case of an addition, an insertion which duplicates some of the text already present.

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **duplicate** all of the text indicated as an addition duplicates some text that is in the original, whether the duplication is word-for-word or less exact.

**duplicate-partial** part of the text indicated as an addition duplicates some text that is in the original

**excessStart** some text at the beginning of the deletion is marked as deleted even though it clearly should not be deleted.

**excessEnd** some text at the end of the deletion is marked as deleted even though it clearly should not be deleted.

**shortStart** some text at the beginning of the deletion is not marked as deleted even though it clearly should be.

**shortEnd** some text at the end of the deletion is not marked as deleted even though it clearly should be.

**partial** some text in the deletion is not marked as deleted even though it clearly should be.

**unremarkable** the deletion is not faulty.[Default]

*Note* Status information on each deletion is needed rather rarely except in critical editions from authorial manuscripts; status information on additions is even less common.

Marking a deletion or addition as faulty is inescapably an interpretive act; the usual test applied in practice is the linguistic acceptability of the text with and without the letters or words in question.

`@cause` documents the presumed cause for the intervention.

*Status* Optional

*Datatype* teidata.enumerated

`@seq` (sequence) assigns a sequence number related to the order in which the encoded features carrying this attribute are believed to have occurred.
*Status* Optional
*Datatype* teidata.count

---

**att.translatable** provides attributes used to indicate the status of a translatable portion of an ODD document.
*Module* tagdocs — Documentation Elements
*Members* desc exemplum gloss remarks valDesc
*Attributes*

`@versionDate` specifies the date on which the source text was extracted and sent to the translator
*Status* Optional
*Datatype* teidata.temporal.working
*Note* The *versionDate* attribute can be used to determine whether a translation might need to be revisited, by comparing the modification date on the containing file with the *versionDate* value on the translation. If the file has changed, changelogs can be checked to see whether the source text has been modified since the translation was made.

---

**att.typed** provides attributes that can be used to classify or subclassify elements in any way. [*1.3.1. Attribute Classes 18.1.1. Words and Above 3.6.1. Referring Strings 3.7. Simple Links and Cross-References 3.6.5. Abbreviations and Their Expansions 3.13.1. Core Tags for Verse 7.2.5. Speech Contents 4.1.1. Un-numbered Divisions 4.1.2. Numbered Divisions 4.2.1. Headings and Trailers 4.4. Virtual Divisions 14.3.2.3. Personal Relationships 12.3.1.1. Core Elements for Transcriptional Work 17.1.1. Pointers and Links 17.3. Blocks, Segments, and Anchors 13.2. Linking the Apparatus to the Text 23.5.1.2. Defining Content Models: RELAX NG 8.3. Elements Unique to Spoken Texts 24.3.1.3. Modification of Attribute and Attribute Value Lists*]
*Module* tei — The TEI Infrastructure
*Members* att.pointing.group[altGrp joinGrp linkGrp] TEI ab abbr accMat add addName addSpan affiliation age alt altIdent altIdentifier am anchor app application bibl biblStruct binaryObject birth bloc c calendar camera castItem cb certainty change cit cl classSpec climate collection colloc constitution constraintSpec corr correspAction correspDesc country custEvent damage damageSpan date death decoNote del delSpan derivation desc dim dimensions distinct district div div1 div2 div3 div4 div5 div6 div7 divGen domain eLeaf eTree education etym event eventName exemplum explicit factuality faith figure filiation finalRubric floatingText forename forest form fw g gb genName gender geogFeat geogName gloss gram gramGrp graph graphic group head iType ident idno incident incipit interaction join kinesic label langKnowledge lb lbl lg line link list listAnnotation listApp listBibl listChange listEvent listForest listNym listObject listOrg listPerson listPlace listRef listRelation location locus m mapping material measure measureGrp media milestone mod moduleSpec move msDesc msFrag msName msPart name nameLink nationality node notatedMusic note noteGrp num nym oRef object objectName occupation offset org orgName origDate origPlace orth path pause pb pc persName persPronouns phr place placeName population post preparedness pron provenance ptr purpose quote rb re recording ref reg region relatedItem relation residence restore rhyme roleName rs rt rubric ruby s schemaRef seal seg settlement sex socecStatus sound spGrp space stamp standOff state surface surfaceGrp surname table tag tech teiCorpus term terrain text time title titlePage titlePart trailer trait unit unitDef usg vocal w witDetail writing xenoData xr zone
*Attributes*

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Status* Optional

*Datatype* teidata.enumerated

```
<div type="verse">
  <head>Night in Tarras</head>
  <lg type="stanza">
    <l>At evening tramping on the hot white road</l>
    <l>…</l>
  </lg>
  <lg type="stanza">
    <l>A wind sprang up from nowhere as the sky</l>
    <l>…</l>
  </lg>
</div>
```

*Note* The *type* attribute is present on a number of elements, not all of which are members of att.typed, usually because these elements restrict the possible values for the attribute in a specific way.

`@subtype` (subtype) provides a sub-categorization of the element, if needed.

*Status* Optional

*Datatype* teidata.enumerated

*Note* The *subtype* attribute may be used to provide any sub-classification for the element additional to that provided by its *type* attribute.

*Schematron* <sch:rule context="tei:*[@subtype]"> <sch:assert test="@type">The <sch:name/> element should not be categorized in detail with @subtype unless also categorized in general with @type</sch:assert> </sch:rule>

*Note* When appropriate, values from an established typology should be used. Alternatively a typology may be defined in the associated TEI header. If values are to be taken from a project-specific list, this should be defined using the <valList> element in the project-specific schema description, as described in *24.3.1.3. Modification of Attribute and Attribute Value Lists* .

---

**att.witnessed** provides attributes used to identify the witnesses supporting a particular reading in a critical apparatus. [*13.1. The Apparatus Entry, Readings, and Witnesses*]

*Module* textcrit — Critical Apparatus

*Members* lem rdg

*Attributes*

`@wit` (witness or witnesses) contains a space-delimited list of one or more pointers indicating the witnesses which attest to a given reading.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Note* If the apparatus contains readings only for a single witness, this attribute may be consistently omitted.

This attribute may occur both within an apparatus gathering variant readings in the transcription of an individual witness and within an apparatus gathering readings from different witnesses.

Additional descriptions or alternative versions of the sigla referenced may be supplied as the content of a child <wit> element.

---

**att.written** provides attributes to indicate the hand in which the content of an element was written in the source being transcribed. [*1.3.1. Attribute Classes*]

*Module* tei — The TEI Infrastructure

*Members* att.damaged[damage damageSpan] att.textCritical[lem rdg rdgGrp] att.transcriptional[add addSpan del delSpan mod redo restore retrace rt subst substJoin undo] ab closer dateline div emph figure fw head hi label line note noteGrp opener p path postscript salute seg signed sp speaker stage text trailer zone

*Attributes*

`@hand` points to a <handNote> element describing the hand considered responsible for the content of the element concerned.

*Status* Optional

*Datatype* teidata.pointer

# C Elements

## C.1 About the Elements Appendix

This appendix gives you links to reference pages for all elements in the TEI Guidelines. There are 588 TEI elements in revision ebf0834e1 of TEI P5 Version 4.10.0 of the TEI Guidelines.

The elements listed here are in the TEI Namespace: http://www.tei-c.org/ns/1.0 unless otherwise noted on that element's reference page.

---

**<ab>** (anonymous block) contains any component-level unit of text, acting as a container for phrase or inter level elements analogous to, but without the same constraints as, a paragraph. [*17.3. Blocks, Segments, and Anchors*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc

  - *@generatedBy*
- att.declaring
  - *@decls*
- att.fragmentable
  - *@part*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*

*Member of* macro.abContent model.pLike

*Contained by*

*cmc:* post

*core:* item note q quote said sp stage textLang

*corpus:* particDesc setting settingDesc

*drama:* castList epilogue performance prologue set view

*figures:* cell figure

*header:* abstract application availability cRefPattern calendar change correction correspAction correspContext correspDesc editionStmt editorialDecl encodingDesc handNote hyphenation interpretation langUsage licence normalization prefixDef projectDesc publicationStmt punctuation quotation refsDecl samplingDecl scriptNote segmentation seriesStmt sourceDesc stdVals styleDefDecl

*linking:* ab

*msdescription:* accMat acquisition additional additions binding bindingDesc collation condition custEvent custodialHist decoDesc decoNote filiation foliation handDesc history layout layoutDesc msContents msDesc msFrag msItem msItemStruct msPart musicNotation objectDesc origin physDesc provenance recordHist scriptDesc seal sealDesc signatures source summary support supportDesc surrogates typeDesc typeNote

*namesdates:* climate event langKnowledge listRelation nym object occupation org person personGrp persona place population state terrain trait

*spoken:* broadcast equipment recording recordingStmt scriptStmt transcriptionDesc

*tagdocs:* exemplum remarks specGrp

*textcrit:* lem rdg

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 epigraph front postscript

*transcr:* metamark

*verse:* metDecl

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The <ab> element may be used at the encoder's discretion to mark any component-level elements in a text for which no other more specific appropriate markup is defined. Unlike paragraphs, <ab> may nest and may use the *type* and *subtype* attributes.

*Example*

```
<div  type="book"  n="Genesis">
  <div  type="chapter"  n="1">
    <ab>In the beginning God created the heaven and the earth.</ab>
    <ab>And the earth was without form, and void; and
        darkness was upon the face of the deep. And the
        spirit of God moved upon the face of the waters.</ab>
    <ab>And God said, Let there be light: and there was light.</ab>
<!-- ...-->
  </div>
</div>
```

*Schematron* <sch:rule context="tei:l//tei:ab"> <sch:assert test="ancestor::tei:floatingText | parent::tei:figure | parent::tei:note"> Abstract model violation: Metrical lines may not contain higher-level divisions such as p or ab, unless ab is a child of figure or note, or is a descendant of floatingText. </sch:assert> </sch:rule>

*Content model*

```
<content>
 <macroRef key="macro.abContent"/>
</content>
```

*Schema Declaration*

```
element ab
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
```

```
   att.cmc.attributes,
   att.declaring.attributes,
   att.fragmentable.attributes,
   att.typed.attributes,
   att.written.attributes,
   macro.abContent}
```

---

**<abbr>** (abbreviation) contains an abbreviation of any sort. [*3.6.5. Abbreviations and Their Expansions*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.typed

- ~~type~~
- @subtype

`@type` (type) allows the encoder to classify the abbreviation according to some convenient typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **suspension** (suspension) the abbreviation provides the first letter(s) of the word or phrase, omitting the remainder.
**contraction** (contraction) the abbreviation omits some letter(s) in the middle.
**brevigraph** the abbreviation comprises a special symbol or mark.
**superscription** (superscription) the abbreviation includes writing above the line.
**acronym** (acronym) the abbreviation comprises the initial letters of the words of a phrase.
**title** (title) the abbreviation is for a title of address (Dr, Ms, Mr, …)
**organization** (organization) the abbreviation is for the name of an organization.
**geographic** (geographic) the abbreviation is for a geographic name.
*Note* The *type* attribute is provided for the sake of those who wish to classify abbreviations at their point of occurrence; this may be useful in some circumstances, though usually the same abbreviation will have the same type in all occurrences. As the sample values make clear, abbreviations may be classified by the method used to construct them, the method of writing them, or the referent of the term abbreviated; the typology used is up to the encoder and should be carefully planned to meet the needs of the expected use. For a typology of Middle English abbreviations, see **?**

*Member of* model.choicePart model.pPart.editorial
*Contained by*
*analysis:* cl pc phr s span w
*cmc:* post
*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit
*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castItem role roleDesc sound tech view
*figures:* cell figDesc
*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage
*iso-fs:* fDescr fsDescr
*linking:* ab seg
*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark
*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* If abbreviations are expanded silently, this practice should be documented in the <editorialDecl>, either with a <normalization> element or a <p>.

*Example*

```
<choice>
  <expan>North Atlantic Treaty Organization</expan>
  <abbr  cert="low">NorATO</abbr>
  <abbr  cert="high">NATO</abbr>
  <abbr  cert="high"  xml:lang="fr">OTAN</abbr>
</choice>
```

*Example*

```
<choice>
  <abbr>SPQR</abbr>
  <expan>senatus populusque romanorum</expan>
</choice>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element abbr
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.phraseSeq}
```

---

**<abstract>** contains a summary or formal abstract prefixed to an existing source document by the encoder.
[*2.4.4. Abstracts*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.profileDescPart

*Contained by*

*header:* profileDesc

*May contain*

*core:* list listBibl p

*figures:* table

*linking:* ab

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*textcrit:* listApp listWit

*Note* This element is intended only for cases where no abstract is available in the original source. Any abstract already present in the source document should be encoded as a  within the <front>, as it should for a born-digital document.

*Example*

```
<profileDesc>
  <abstract  resp="#LB">
    <p>Good database design involves the acquisition and deployment of
        skills which have a wider relevance to the educational process. From
        a set of more or less instinctive rules of thumb a formal discipline
        or "methodology" of database design has evolved. Applying that
        methodology can be of great benefit to a very wide range of academic
        subjects: it requires fundamental skills of abstraction and
        generalisation and it provides a simple mechanism whereby complex
        ideas and information structures can be represented and manipulated,
        even without the use of a computer. </p>
  </abstract>
</profileDesc>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.pLike"/>
  <classRef key="model.listLike"/>
  <elementRef key="listBibl"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element abstract
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.pLike | model.listLike | listBibl )+
}
```

---

**<accMat>** (accompanying material) contains details of any significant additional material which may be closely associated with the manuscript or object being described, such as non-contemporaneous documents or fragments bound in with it at some earlier historical period. [*11.7.3.3. Accompanying Material*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.physDescPart

*Contained by*

*msdescription:* physDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<accMat>A copy of a tax form from 1947 is included in the envelope
with the letter. It is not catalogued separately.</accMat>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
```

```
</content>
```

*Schema Declaration*

```
element accMat
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   macro.specialPara}
```

---

**<acquisition>** (acquisition) contains any descriptive or other information concerning the process by which a manuscript or manuscript part or other object entered the holding institution. [*11.8. History*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*

    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*

*Contained by:* *msDescription:* history
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound spGrp tech view
*figures:* figure formula notatedMusic table
*gaiji:* g
*header:* biblFull idno
*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<acquisition>Left to the <name  type="place">Bodleian</name> by
<name  type="person">Richard Rawlinson</name> in 1755.
</acquisition>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element acquisition
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   macro.specialPara}
```

---

## <activity>

(activity) contains a brief informal description of what a participant in a language interaction is doing other than speaking, if anything. [*16.2.3. The Setting Description*]

*Module* corpus — Language Corpora
*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.settingPart
*Contained by*
*corpus:* setting
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit
*dictionaries:* lang
*figures:* figure notatedMusic
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data

*Note* For more fine-grained description of participant activities during a spoken text, the <event> element should be used.

*Example*

```
<activity>driving</activity>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element activity
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq.limited}
```

---

## <actor> contains the name of an actor appearing within a cast list. [*7.1.4. Cast Lists*]

*Module* drama — Performance Texts

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.canonical
  - *@key*
  - *@ref*

`@sex` specifies the sex of the actor.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.sex separated by whitespace
*Note* Values for this attribute may be locally defined by a project, or may refer to an external standard.

`@gender` specifies the gender of the actor.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.gender separated by whitespace
*Note* Values for this attribute may be locally defined by a project, or they may refer to an external standard.

*Member of* model.castItemPart

*Contained by*

*drama:* castItem

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* This element should be used only to mark the name of the actor as given in the source. Chapter *14. Names, Dates, People, and Places* discusses ways of marking the components of names, and also of associating names with biographical information about a person.

*Example*

```
<castItem>
  <role>Mathias</role>
  <roleDesc>the Burgomaster</roleDesc>
  <actor  ref="https://en.wikipedia.org/wiki/Henry_Irving">Mr. Henry
Irving</actor>
</castItem>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element actor
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   attribute sex { list { + } }?,
   attribute gender { list { + } }?,
   macro.phraseSeq}
```

**<add>** (addition) contains letters, words, or phrases inserted in the source text by an author, scribe, or a previous annotator or corrector. [*3.5.3. Additions, Deletions, and Omissions*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*
  - att.ranging

    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
  - att.transcriptional
    - *@status*
    - *@cause*
    - *@seq*
    - att.editLike
    * *@evidence*
    * *@instant*
    - att.placement
    * *@place*
    - att.written
    * *@hand*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.linePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw line metamark mod restore retrace secl subst supplied surplus zone

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* In a diplomatic edition attempting to represent an original source, the <add> element should not be used for additions to the current TEI electronic edition made by editors or encoders. In these cases, either the <corr> or <supplied> element are recommended.

In a TEI edition of a historical text with previous editorial emendations in which such additions or reconstructions are considered part of the source text, the use of <add> may be appropriate, dependent on the editorial philosophy of the project.

*Example*

```
The story I am
going to relate is true as to its main facts, and as to the
consequences <add  place="above">of these facts</add> from which
this tale takes its title.
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element add
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.transcriptional.attributes,
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   att.typed.attributes,
   macro.paraContent}
```

---

**<additional>** (additional) groups additional information, combining bibliographic information about a manuscript or other object, or surrogate copies of it, with curatorial or administrative information. [*11.9. Additional Information*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Contained by:* *msDescription:* msDesc msFrag msPart
*namesdates:* object
*May contain*
*core:* listBibl p
*linking:* ab
*msdescription:* adminInfo surrogates
*Example*

```
<additional>
  <adminInfo>
    <recordHist>
      <p>
<!-- record history here -->
      </p>
    </recordHist>
    <custodialHist>
      <p>
<!-- custodial history here -->
      </p>
    </custodialHist>
  </adminInfo>
  <surrogates>
    <p>
<!-- information about surrogates here -->
    </p>
  </surrogates>
  <listBibl>
    <bibl>
<!-- ... -->
    </bibl>
<!-- full bibliography here -->
  </listBibl>
</additional>
```

*Content model*

```
<content>
 <alternate>
  <sequence>
   <elementRef key="adminInfo"
    minOccurs="0"/>
   <elementRef key="surrogates"
    minOccurs="0"/>
   <elementRef key="listBibl" minOccurs="0"/>
```

```
  </sequence>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element additional
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( ( adminInfo?, surrogates?, listBibl? ) | model.pLike+ )
}
```

---

## <additions>

<additions> (additions) contains a description of any significant additions found within a manuscript or other object, such as marginalia or other annotations. [*11.7.2. Writing, Decoration, and Other Notations*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.physDescPart

*Contained by*

*msdescription:* physDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<additions>
  <p>There are several marginalia in this manuscript. Some consist of
     single characters and others are figurative. On 8v is to be found a
```

```
drawing of
    a mans head wearing a hat. At times sentences occurs: On 5v:
  <q  xml:lang="is">Her er skrif andres isslendin</q>,
    on 19r: <q  xml:lang="is">þeim go</q>,
    on 21r: <q  xml:lang="is">amen med aund ok munn halla rei knar hofud summu
all huad
        batar þad mælgi ok mal</q>,
    On 21v: some runic letters and the sentence <q  xml:lang="la">aue maria
gracia plena dominus</q>.</p>
</additions>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element additions
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.specialPara}
```

---

**<addName>** (additional name) contains an additional name component, such as a nickname, epithet, or alias, or any other descriptive phrase used within a personal name. [*14.2.1. Personal Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.personal
  - *@full*
  - *@sort*
  - att.naming
    * *@role*
    * *@nymRef*
    * att.canonical
      · *@key*
      · *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.persNamePart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<persName>
  <forename>Frederick</forename>
  <addName  type="epithet">the Great</addName>
  <roleName>Emperor of Prussia</roleName>
</persName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element addName
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.personal.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

## <address>

**<address>** (address) contains a postal address, for example of a publisher, an organization, or an individual. [*3.6.2. Addresses 2.2.4. Publication, Distribution, Licensing, etc. 3.12.2.4. Imprint, Size of a Document, and Reprint Information*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*

*Member of* model.addressLike model.publicationStmtPart.detail

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal publicationStmt rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectIdentifier objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* addrLine cb ellipsis gap gb index lb milestone name note noteGrp pb postBox postCode rs street

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*namesdates:* addName bloc climate country district eventName forename genName geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* This element should be used for postal addresses only. Within it, the generic element <addrLine> may be used as an alternative to any of the more specialized elements available from the `model.addrPart` class, such as <street>, <postCode> etc.

*Example* Using just the elements defined by the core module, an address could be represented as follows:

```
<address>
  <street>via Marsala 24</street>
  <postCode>40126</postCode>
  <name>Bologna</name>
  <name>Italy</name>
</address>
```

*Example* When a schema includes the names and dates module more specific elements such as country or settlement would be preferable over generic <name>:

```
<address>
  <street>via Marsala 24</street>
  <postCode>40126</postCode>
  <settlement>Bologna</settlement>
  <country>Italy</country>
</address>
```

*Example*

```
<address>
  <addrLine>Computing Center, MC 135</addrLine>
  <addrLine>P.O. Box 6998</addrLine>
  <addrLine>Chicago, IL 60680</addrLine>
  <addrLine>USA</addrLine>
</address>
```

*Example*

```
<address>
  <country key="FR"/>
  <settlement type="city">Lyon</settlement>
  <postCode>69002</postCode>
  <district type="arrondissement">IIème</district>
  <district type="quartier">Perrache</district>
  <street>
    <num>30</num>, Cours de Verdun</street>
</address>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.global"
   minOccurs="0" maxOccurs="unbounded"/>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.addrPart"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element address
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   ( model.global*, ( ( model.addrPart, model.global* )+ ) )
}
```

---

**<addrLine>** (address line) contains one line of a postal address. [*3.6.2. Addresses 2.2.4. Publication, Distribution, Licensing, etc. 3.12.2.4. Imprint, Size of a Document, and Reprint Information*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.addrPart

*Contained by*

*core:* address

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Addresses may be encoded either as a sequence of lines, or using any sequence of component elements from the `model.addrPart` class. Other non-postal forms of address, such as telephone numbers or email, should not be included within an <address> element directly but may be wrapped within an <addrLine> if they form part of the printed address in some source text.

*Example*

```
<address>
  <addrLine>Computing Center, MC 135</addrLine>
  <addrLine>P.O. Box 6998</addrLine>
  <addrLine>Chicago, IL</addrLine>
  <addrLine>60680 USA</addrLine>
</address>
```

*Example*

```
<addrLine>
  <ref  target="tel:+1-201-555-0123">(201) 555 0123</ref>
</addrLine>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element addrLine
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

---

## <addSpan>

(added span of text) marks the beginning of a longer sequence of text added by an author, scribe, annotator or corrector (see also <add>). [*12.3.1.4. Additions and Deletions*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*

  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.spanning
  - *@spanTo*

- att.transcriptional
  - *@status*
  - *@cause*
  - *@seq*

  - att.editLike
    - *@evidence*
    - *@instant*
  - att.placement
    - *@place*
  - att.written
    - *@hand*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.global.edit

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Note* Both the beginning and the end of the added material must be marked; the beginning by the <addSpan> element itself, the end by the *spanTo* attribute.

*Example*

```
<handNote  xml:id="HEOL"
  scribe="HelgiÓlafsson"/>
<!-- ... -->
<body>
  <div>
<!-- text here -->
  </div>
  <addSpan  n="added_gathering"  hand="#HEOL"
    spanTo="#P025"/>
  <div>
<!-- text of first added poem here -->
  </div>
  <div>
<!-- text of second added poem here -->
  </div>
  <div>
<!-- text of third added poem here -->
  </div>
  <div>
<!-- text of fourth added poem here -->
  </div>
  <anchor  xml:id="P025"/>
  <div>
<!-- more text here -->
  </div>
</body>
```

*Schematron* <sch:rule context="tei:addSpan"> <sch:assert test="@spanTo">The @spanTo attribute of <sch:name/> is required.</sch:assert> </sch:rule>

*Schematron* <sch:rule context="tei:addSpan"> <sch:assert test="@spanTo">L'attribut spanTo est requis.</sch:assert> </sch:rule>

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element addSpan
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.spanning.attributes,
   att.transcriptional.attributes,
```

```
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   att.typed.attributes,
   empty
}
```

---

**<adminInfo>** (administrative information) contains information about the present custody and availability of the manuscript or other object, and also about the record description itself. [*11.9.1. Administrative Information*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by:* *msdescription:* additional
*May contain*
*core:* note noteGrp
*header:* availability
*msdescription:* custodialHist recordHist
*Example*

```
<adminInfo>
  <recordHist>
    <source>Record created <date>1 Aug 2004</date>
    </source>
  </recordHist>
  <availability>
    <p>Until 2015 permission to photocopy some materials from this
        collection has been limited at the request of the donor. Please ask
repository staff for details
        if you are interested in obtaining photocopies from Series 1:
        Correspondence.</p>
  </availability>
  <custodialHist>
    <p>Collection donated to the Manuscript Library by the Estate of
        Edgar Holden in 1993. Donor number: 1993-034.</p>
  </custodialHist>
</adminInfo>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="recordHist"
   minOccurs="0"/>
  <elementRef key="availability"
   minOccurs="0"/>
  <elementRef key="custodialHist"
   minOccurs="0"/>
  <classRef key="model.noteLike"
   minOccurs="0"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element adminInfo
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( recordHist?, availability?, custodialHist?, model.noteLike? )
}
```

**<affiliation>** (affiliation) contains an informal description of a person's present or past affiliation with some organization, for example an employer or sponsor. [*16.2.2. The Participant Description*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*

  - *@to-custom*
  - *@datingPoint*
  - *@datingMethod*
- att.datable.iso
  - *@when-iso*
  - *@notBefore-iso*
  - *@notAfter-iso*
  - *@from-iso*
  - *@to-iso*
- att.datable.w3c
  - *@when*
  - *@notBefore*
  - *@notAfter*
  - *@from*
  - *@to*

- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*

- att.canonical
  - *@key*
  - *@ref*
- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **sponsor**
**recommend**
**discredit**
**pledged**

*Member of* model.addressLike model.persStateLike
*Contained by*
*analysis:* cl phr s span
*cmc:* post
*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p

pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* If included, the name of an organization may be tagged using either the <name> element as above, or the more specific <orgName> element.

*Example*

```
<affiliation>Junior project officer for the US <name  type="org">National
Endowment for
    the Humanities</name>
</affiliation>
```

*Example* This example indicates that the person was affiliated with the Australian Journalists Association at some point between the dates listed.

```
<affiliation  notAfter="1960-01-01"
  notBefore="1957-02-28">Paid up member of the
<orgName>Australian Journalists Association</orgName>
</affiliation>
```

*Example* This example indicates that the person was affiliated with Mount Holyoke College throughout the entire span of the date range listed.

```
<affiliation  from="1902-01-01"
  to="1906-01-01">Was an assistant professor at Mount Holyoke
College.</affiliation>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element affiliation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attribute.subtype,
```

```
   attribute type { text }?,
   macro.phraseSeq}
```

---

**<age>** (age) specifies the age of a person. [*14.3.2.1. Personal Characteristics*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*

  * *@notAfter-custom*
  * *@from-custom*
  * *@to-custom*
  * *@datingPoint*
  * *@datingMethod*
- att.datable.iso
  * *@when-iso*
  * *@notBefore-iso*
  * *@notAfter-iso*
  * *@from-iso*
  * *@to-iso*
- att.datable.w3c
  * *@when*
  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*

- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.editLike
  - *@evidence*
  - *@instant*
- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **western**
**sui**

**subjective**

**objective**

**inWorld** (in world) age of a fictional character at the time the story takes place, rather than at the time the story is told

**chronological**

**biological**

**psychological**

**functional**

`@value` supplies a numeric code representing the age or age group.

*Status* Optional

*Datatype* teidata.count

*Note* This attribute may be used to complement a more detailed discussion of a person's age in the content of the element

*Member of* model.persStateLike

*Contained by*

*namesdates:* person personGrp persona

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin

character data

*Note* As with other culturally-constructed traits such as sex, the way in which this concept is described in different cultural contexts may vary. The normalizing attributes are provided as a means of simplifying that variety to Western European norms and should not be used where that is inappropriate. The content of the element may be used to describe the intended concept in more detail, using plain text.

*Example*

```
<age value="2" notAfter="1986">under 20 in the early eighties</age>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
```

```
</content>
```

*Schema Declaration*

```
element age
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   attribute value { text }?,
   macro.phraseSeq.limited}
```

---

**<alt>** (alternation) identifies an alternation or a set of choices among elements or passages. [*17.8. Alternation*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

- att.cmc
  - *@generatedBy*
- att.typed
  - *@type*
  - *@subtype*
- att.pointing
  - ~~target~~
  - @targetLang
  - @evaluate

`@target` specifies the destination of the reference by supplying one or more URI References.
*Derived from* att.pointing
*Status* Optional
*Datatype* 2–∞ occurrences of teidata.pointer separated by whitespace

`@mode` states whether the alternations gathered in this collection are exclusive or inclusive.
*Status* Recommended
*Datatype* teidata.enumerated
*Legal values are:* **excl** (exclusive) indicates that the alternation is exclusive, i.e. that at most one of the alternatives occurs.
**incl** (inclusive) indicates that the alternation is not exclusive, i.e. that one or more of the alternatives occur.

`@weights` If *mode* is excl, each weight states the probability that the corresponding alternative occurs. If *mode* is incl each weight states the probability that the corresponding alternative occurs given that at least one of the other alternatives occurs.
*Status* Optional
*Datatype* 2–∞ occurrences of teidata.probability separated by whitespace
*Note* If *mode* is excl, the sum of weights must be 1. If *mode* is incl, the sum of weights must be in the range from 0 to the number of alternants.

*Member of* model.global.meta
*Contained by*
*analysis:* cl m phr s span w
*cmc:* post
*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note

num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab altGrp seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Example*

```
<alt  mode="excl"  target="#we.fun #we.sun"
  weights="0.5 0.5"/>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element alt
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.typed.attributes,
   att.pointing.attribute.targetLang,
   att.pointing.attribute.evaluate,
   attribute target { list { * } }?,
```

```
   attribute mode { "excl" | "incl" }?,
   attribute weights { list { * } }?,
   empty
}
```

---

**<alternate>** indicates that the constructs referenced by its children form an alternation. [*23.5.1. Defining Content Models*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.repeatable
  - *@minOccurs*
  - *@maxOccurs*

*Member of* model.contentPart

*Contained by*

*tagdocs:* alternate content sequence

*May contain*

*tagdocs:* alternate anyElement classRef dataRef elementRef empty macroRef sequence textNode valList

*Example*

```
<content>
  <alternate>
    <elementRef  key="name"/>
    <elementRef  key="persName"/>
  </alternate>
</content>
```

This example content model permits either a <name> or a <persName>.

*Schematron* <sch:rule context="tei:alternate"> <sch:assert test="count(*) gt 1">The alternate element must have at least two child elements</sch:assert> </sch:rule>

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <elementRef key="valList"/>
  <classRef key="model.contentPart"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element alternate
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.repeatable.attributes,
   ( valList | model.contentPart )+
}
```

---

## <altGrp>

(alternation group) groups a collection of <alt> elements and possibly pointers. [*17.8. Alternation*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.pointing.group
  - *@domains*
  - *@targFunc*

  - att.pointing
    - *@targetLang*
    - *@target*
    - *@evaluate*
  - att.typed
    - *@type*
    - *@subtype*

`@mode` states whether the alternations gathered in this collection are exclusive or inclusive.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **excl** (exclusive) indicates that the alternation is exclusive, i.e. that at most one of the alternatives occurs.[Default]

**incl** (inclusive) indicates that the alternation is not exclusive, i.e. that one or more of the alternatives occur.

*Member of* model.global.meta

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*core:* desc ptr

*linking:* alt

*Note* Any number of alternations, pointers or extended pointers.

*Example*

```
<altGrp mode="excl">
  <alt target="#dm #lt #bb"
    weights="0.5 0.25 0.25"/>
  <alt target="#rl #db" weights="0.5 0.5"/>
</altGrp>
```

*Example*

```
<altGrp  mode="incl">
  <alt  target="#dm #rl"  weights="0.90 0.90"/>
  <alt  target="#lt #rl"  weights="0.5 0.5"/>
  <alt  target="#bb #rl"  weights="0.5 0.5"/>
  <alt  target="#dm #db"  weights="0.10 0.10"/>
  <alt  target="#lt #db"  weights="0.45 0.90"/>
  <alt  target="#bb #db"  weights="0.45 0.90"/>
</altGrp>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.descLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="alt"/>
   <elementRef key="ptr"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element altGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.pointing.group.attributes,
   att.pointing.attributes,
   att.typed.attributes,
   attribute mode { "excl" | "incl" }?,
   ( model.descLike*, ( alt | ptr )* )
}
```

---

**<altIdent>** (alternate identifier) supplies the recommended XML name for an element, class, attribute, etc. in some language. [*3.4.1. Terms and Glosses 23.4.1. Description of Components*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

- att.typed
  - *@type*
  - *@subtype*

*Member of* model.identSynonyms

*Contained by*

*tagdocs:* attDef classSpec constraintSpec dataSpec elementSpec macroSpec valItem

*May contain* XSD NCName

*Note* All documentation elements in ODD have a canonical name, supplied as the value for their *ident* attribute. The <altIdent> element is used to supply an alternative name for the corresponding XML object, perhaps in a different language.

*Example*

```
<altIdent xml:lang="fr">balisageDoc</altIdent>
```

*Content model* `<content> <dataRef name="NCName"/></content>`

*Schema Declaration*

```
element altIdent
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   xsd:NCName
}
```

---

**<altIdentifier>** (alternative identifier) contains an alternative or former structured identifier used for a manuscript or other object, such as a former catalogue number. [*11.4. The Manuscript Identifier*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.typed
  - *@type*
  - *@subtype*

*Contained by:* *msDescription:* msFrag msIdentifier
*namesdates:* objectIdentifier
*May contain*
*core:* note
*header:* idno
*msdescription:* collection institution repository
*namesdates:* bloc country district geogName placeName region settlement
*Note* An identifying number of some kind must be supplied if known; if it is not known, this should be stated.
*Example*

```
<altIdentifier>
  <settlement>San Marino</settlement>
  <repository>Huntington Library</repository>
  <idno>MS.El.26.C.9</idno>
</altIdentifier>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.placeNamePart"
   expand="sequenceOptional"/>
  <elementRef key="institution"
   minOccurs="0"/>
  <elementRef key="repository"
   minOccurs="0"/>
  <elementRef key="collection"
   minOccurs="0"/>
  <elementRef key="idno"/>
  <elementRef key="note" minOccurs="0"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element altIdentifier
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.typed.attributes,
   (
      placeName?,
      bloc?,
      country?,
      region?,
      settlement?,
      district?,
      geogName?,
      institution?,
      repository?,
      collection?,
      idno,
      note?
   )
}
```

---

**<am>** (abbreviation marker) contains a sequence of letters or signs present in an abbreviation which are omitted or replaced in the expanded form of the abbreviation. [*12.3.1.2. Abbreviation and Expansion*]

*Module* transcr — Representation of Primary Sources

*Attributes*
- att.global
  - *@xml:id*

  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.editLike
  - *@evidence*
  - *@instant*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.choicePart model.pPart.editorial

*Contained by*

*analysis:* cl pc phr s span w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*core:* add corr del orig reg sic unclear

*gaiji:* g

*transcr:* damage handShift mod redo restore retrace secl supplied surplus undo

character data

*Example*

```
 do you <abbr>Mr<am>.</am>
</abbr> Jones?
```

*Example*

```
<choice>
  <abbr>Aug<am>g</am>
  </abbr>
  <expan>Aug<ex>ustorum duo</ex>
  </expan>
</choice>
```

*Example*

```
<abbr>eu<am>
    <g  ref="#b-er"/>
  </am>y</abbr>
<abbr>
  <am>
    <g  ref="#b-per"/>
  </am>sone
```

```
</abbr> ...
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.pPart.transcriptional"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element am
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.editLike.attributes,
   att.typed.attributes,
   ( text | model.gLike | model.pPart.transcriptional )*
}
```

---

## <analytic>

<analytic> (analytic level) contains bibliographic elements describing an item (e.g. an article or poem) published within a monograph or journal and not as an independent publication. [*3.12.2.1. Analytic, Monographic, and Series Levels*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Contained by* biblStruct

*May contain*

*core:* author date editor ptr ref respStmt textLang title

*header:* availability idno

*tagdocs:* listRef

*Note* May contain titles and statements of responsibility (author, editor, or other), in any order.

The <analytic> element may only occur within a <biblStruct>, where its use is mandatory for the description of an analytic level bibliographic item.

*Example*

```
<biblStruct>
  <analytic>
    <author>Chesnutt, David</author>
    <title>Historical Editions in the States</title>
  </analytic>
  <monogr>
    <title  level="j">Computers and the Humanities</title>
    <imprint>
      <date  when="1991-12">(December, 1991):</date>
    </imprint>
    <biblScope>25.6</biblScope>
    <biblScope>377–380</biblScope>
  </monogr>
</biblStruct>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <elementRef key="author"/>
  <elementRef key="editor"/>
```

```
  <elementRef key="respStmt"/>
  <elementRef key="title"/>
  <classRef key="model.ptrLike"/>
  <elementRef key="date"/>
  <elementRef key="textLang"/>
  <elementRef key="idno"/>
  <elementRef key="availability"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element analytic
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      author    | editor    | respStmt    | title    | model.ptrLike    | date    | textLa
}
```

---

## <anchor>

(anchor point) attaches an identifier to a point within a text, whether or not it corresponds with a textual element. [*8.4.2. Synchronization and Overlap 17.5. Correspondence and Alignment*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.milestoneLike

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list listBibl measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor

docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc subst supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Note* On this element, the global *xml:id* attribute must be supplied to specify an identifier for the point at which this element occurs within a document. The value used may be chosen freely provided that it is unique within the document and is a syntactically valid name. There is no requirement for values containing numbers to be in sequence.

*Example*

```
<s>The anchor is he<anchor  xml:id="A234"/>re somewhere.</s>
<s>Help me find it.<ptr  target="#A234"/>
</s>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element anchor
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.typed.attributes,
   empty
}
```

---

## <annotation>

**<annotation>** represents an annotation following the Web Annotation Data Model. [*17.10. The standOff Container*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*

- att.global
  - ~~xml:id~~
  - @n
  - @xml:lang
  - @xml:base
  - @xml:space
- att.global.analytic
  - @ana
- att.global.change
  - @change
- att.global.facs
  - @facs

- att.global.linking
  - @corresp
  - @synch
  - @sameAs
  - @copyOf
  - @next
  - @prev
  - @exclude
  - @select
- att.global.rendition
  - @rend
  - @style
  - @rendition
- att.global.responsibility
  - @cert
  - @resp
- att.global.source
  - @source
- att.pointing
  - ~~target~~
  - @targetLang
  - @evaluate

`@xml:id` (identifier) provides a unique identifier for the element bearing the attribute.
*Derived from* att.global
*Status* Required
*Datatype* ID

`@target` specifies the destination of the reference by supplying one or more URI References.
*Derived from* att.pointing
*Status* Required
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

`@motivation` *Status* Optional
*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace
*Legal values are:* **assessing** intent is to assess the target resource in some way, rather than simply make a comment about it
**bookmarking** intent is to create a bookmark to the target or part thereof
**classifying** intent is to classify the target in some way
**commenting** intent is to comment about the target
**describing** intent is to describe the target, rather than (for example) comment on it
**editing** intent is to request an edit or a change to the target resource
**highlighting** intent is to highlight the target resource or a segment thereof
**identifying** intent is to assign an identity to the target
**linking** intent is to link to a resource related to the target
**moderating** intent is to assign some value or quality to the target
**questioning** intent is to ask a question about the target

**replying** intent is to reply to a previous statement, either an annotation or another resource

**tagging** intent is to associate a tag with the target

*Note* For further detailed explanation of the suggested values, see the Web Annotation Vocabulary (WAV). The motivations described here map to URIs defined by the WAV and when exported to RDF or JSON-LD must have the URI `http://www.w3.org/ns/oa#` prepended.

As an RDF vocabulary, WADM permits the definition of new motivations (see Appendix C of the WAV). In TEI, new motivations may be defined in a custom ODD (see section 23.3.1.3). New motivations must also map to URIs defined by an RDF ontology extending the WAV.

*Member of* model.annotationLike

*Contained by*

*linking:* listAnnotation standOff

*May contain*

*core:* note ptr ref respStmt

*header:* licence revisionDesc

*Example*

```
<annotation xml:id="ann1"
  motivation="linking" target="#Gallia">
<!-- See https://www.w3.org/TR/annotation-model/#lifecycle-information and
https://www.w3.org/TR/annotation-model/#agents -->
  <respStmt xml:id="fred">
    <resp>creator</resp>
    <persName>Fred Editor</persName>
  </respStmt>
  <revisionDesc>
    <change status="created"
      when="2020-05-21T13:59:00Z" who="#fred"/>
    <change status="modified"
      when="2020-05-21T19:48:00Z" who="#fred"/>
  </revisionDesc>
<!-- See https://www.w3.org/TR/annotation-model/#rights-information -->
  <licence target="http://creativecommons.org/licenses/by/3.0/"/>
<!-- Multiple bodies -->
<!-- Pointers to sections of text in the same document -->
  <ptr target="#string-range(c1p1s1,0,6)"/>
  <ptr target="#string-range(c1p1s6,19,7)"/>
</annotation>
```

*Example*

```
<annotation xml:id="TheCorrectTitle"
  motivation="commenting" target="#line1">
  <note>The correct title of this specification, and the correct full name of
XML, is
    "Extensible Markup Language". "eXtensible Markup Language" is just a
spelling error.
    However, the abbreviation "XML" is not only correct but, appearing as it
does in the title
    of the specification, an official name of the Extensible Markup Language.
</note>
</annotation>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="respStmt" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="revisionDesc"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="licence" minOccurs="0"
   maxOccurs="unbounded"/>
  <classRef key="model.annotationPart.body"
   minOccurs="0" maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element annotation
{
   att.global.attribute.n,
   att.global.attribute.xmllang,
   att.global.attribute.xmlbase,
   att.global.attribute.xmlspace,
   att.global.analytic.attribute.ana,
   att.global.change.attribute.change,
   att.global.facs.attribute.facs,
   att.global.linking.attribute.corresp,
   att.global.linking.attribute.synch,
   att.global.linking.attribute.sameAs,
   att.global.linking.attribute.copyOf,
   att.global.linking.attribute.next,
   att.global.linking.attribute.prev,
   att.global.linking.attribute.exclude,
   att.global.linking.attribute.select,
   att.global.rendition.attribute.rend,
   att.global.rendition.attribute.style,
   att.global.rendition.attribute.rendition,
   att.global.responsibility.attribute.cert,
   att.global.responsibility.attribute.resp,
   att.global.source.attribute.source,
   att.pointing.attribute.targetLang,
   att.pointing.attribute.evaluate,
   attribute xml:id { text },
   attribute target { list { + } },
   attribute motivation
   {
      list
      {
         (
            "assessing"
          | "bookmarking"
          | "classifying"
          | "commenting"
          | "describing"
          | "editing"
          | "highlighting"
          | "identifying"
          | "linking"
```

```
          | "moderating"
          | "questioning"
          | "replying"
          | "tagging"
         )+
      }
   }?,
   ( respStmt*, revisionDesc*, licence*, model.annotationPart.body* )
}
```

---

**<annotationBlock>** groups together various annotations, e.g. for parallel interpretations of a spoken segment. [*8.4.6. Analytic Coding*]

*Module* spoken — Transcriptions of Speech

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.ascribed
  - *@who*
- att.timed
  - *@start*
  - *@end*

  - att.duration
  - att.duration.iso
    - *@dur-iso*
  - att.duration.w3c
    - *@dur*

*Member of* model.annotationLike model.divPart.spoken
*Contained by*
*core:* item note q quote said stage textLang
*drama:* castList epilogue performance prologue set view
*figures:* cell figure
*header:* change handNote licence scriptNote
*linking:* listAnnotation standOff
*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote
*namesdates:* occupation
*tagdocs:* specGrp
*textcrit:* lem rdg
*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript
*transcr:* metamark
*May contain*
*analysis:* spanGrp
*spoken:* incident kinesic pause shift u vocal writing
*Example*

```
<annotationBlock who="#SPK1" start="#T2"
  end="#T3" xml:id="ag20">
  <u xml:id="u20">
    <seg xml:id="seg37" type="utterance"
      subtype="modeless">
      <w xml:id="w46">Yeah</w>
    </seg>
  </u>
</annotationBlock>
<annotationBlock who="#SPK1" start="#T5"
  end="#T6" xml:id="ag21">
  <u xml:id="u21">
    <seg xml:id="seg38" type="utterance"
      subtype="modeless">
      <w xml:id="w47">Mhm</w>
    </seg>
  </u>
</annotationBlock>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <elementRef key="u"/>
  <elementRef key="spanGrp"/>
  <classRef key="model.global.spoken"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element annotationBlock
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.attributes,
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   ( u | spanGrp | model.global.spoken )*
}
```

---

**<anyElement>** indicates the presence of any elements in a content model. [*23. Documentation Elements*]

*Module* tagdocs — Documentation Elements

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*

    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.repeatable
  - *@minOccurs*
  - *@maxOccurs*

`@require` supplies a list of namespaces to one of which the permitted elements must belong.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.namespace separated by whitespace

`@except` supplies a list of namespaces or prefixed element names which are not permitted.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.namespaceOrName separated by whitespace

*Note* When RELAX NG DTD Compatibility Mode is turned on, validation requires that any elements which may take an *xml:id* be excluded from the content of `anyName` name classes. For normal TEI purposes, this includes anything in the TEI namespace and the <egXML> element. Default elements and namespaces to be excluded are set on the *defaultExceptions* attribute of the <schemaSpec>. *except* may be used when the exceptions need to be locally specified. Element prefixes must be declared using a namespace declaration on the parent or an ancestor element.

*Member of* model.contentPart

*Contained by*

*tagdocs:* alternate content sequence

*May contain* Empty element

*Example*

```
<content>
  <anyElement  except="http://www.tei-c.org/ns/1.0"/>
</content>
```

This <content> element permits one occurrence of any element which does not belong to the TEI namespace.

*Example*

```
<content>
  <anyElement  minOccurs="1"
    maxOccurs="unbounded"
    require="http://www.w3.org/1998/Math/MathML
```

```
http://www.tei-c.org/ns/Examples/1.0"/>
</content>
```

This <content> element permits one or more elements which belong either to the MathML namespace or to the TEI examples namespace.

*Example*

```
<content>
  <alternate  minOccurs="1"
    maxOccurs="unbounded">
    <textNode/>
    <anyElement  require="http://www.tei-c.org/ns/Examples/1.0"/>
  </alternate>
</content>
```

This <content> element permits one or more children which are either text nodes or elements from the TEI examples namespace. It is the equivalent of the content model ANY in an XML DTD using the TEI All schema.

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element anyElement
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.repeatable.attributes,
   attribute require { list { + } }?,
   attribute except { list { + } }?,
   empty
}
```

---

**<app>** (apparatus entry) contains one entry in a critical apparatus, with an optional lemma and usually one or more readings or notes on the relevant passage. [*13.1.1. The Apparatus Entry*]

*Module* textcrit — Critical Apparatus

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.typed
  - ~~type~~
  - @subtype

`@type` classifies the variation contained in this element according to some convenient typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated

`@from` identifies the beginning of the lemma in the base text.
*Status* Optional
*Datatype* teidata.pointer
*Note* This attribute should be used when either the double-end point method of apparatus markup, or the location-referenced method with a URL rather than canonical reference, are used.

`@to` identifies the endpoint of the lemma in the base text.
*Status* Optional
*Datatype* teidata.pointer
*Note* This attribute is only used when the double-end point method of apparatus markup is used, when the encoded apparatus is not embedded *in-line* in the base-text.

`@loc` (location) indicates the location of the variation, when the location-referenced method of apparatus markup is used.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.word separated by whitespace
*Note* This attribute is used only when the location-referenced encoding method is used. It supplies a string containing a canonical reference for the passage to which the variation applies.

*Member of* model.global.edit

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem listApp rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*core:* note noteGrp

*textcrit:* lem rdg rdgGrp wit witDetail

*Example*

```
<app>
  <lem  wit="#El #Hg">Experience</lem>
  <rdg  wit="#La"  type="substantive">Experiment</rdg>
  <rdg  wit="#Ra2"  type="substantive">Eryment</rdg>
</app>
```

*Example*

```
<app  type="substantive">
  <rdgGrp  type="subvariants">
    <lem  wit="#El #Hg">Experience</lem>
    <rdg  wit="#Ha4">Experiens</rdg>
```

```
    </rdgGrp>
    <rdgGrp type="subvariants">
      <lem wit="#Cp #Ld1">Experiment</lem>
      <rdg wit="#La">Ex<g ref="#per"/>iment</rdg>
    </rdgGrp>
    <rdgGrp type="subvariants">
      <lem resp="#ed2013">Eriment</lem>
      <rdg wit="#Ra2">Eryment</rdg>
    </rdgGrp>
  </app>
```

*Example*

```
  <app loc="1">
    <rdg resp="#SEG">ΤΙΜΩΔΑ</rdg>
  </app>
```

*Example*

```
  <app loc="1-6">
    <note>Too badly worn to yield a text</note>
  </app>
```

*Example*

```
  <choice xml:id="choice3">
    <reg>σύμπαντα</reg>
    <orig>ΣΙΝΠΑΤΑΝ</orig>
  </choice>
  <!-- ... -->
  <app from="#choice3">
    <note>Mommsen's fanciful normalization, reproduced here, has not been
  accepted by all recent editions</note>
  </app>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="lem" minOccurs="0"/>
  <alternate maxOccurs="unbounded"
   minOccurs="0">
   <classRef key="model.rdgLike"/>
   <classRef key="model.noteLike"/>
   <elementRef key="witDetail"/>
   <elementRef key="wit"/>
   <elementRef key="rdgGrp"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element app
{
   att.global.attributes,
   att.global.analytic.attributes,
```

```
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   attribute from { text }?,
   attribute to { text }?,
   attribute loc { list { + } }?,
   ( lem?, ( model.rdgLike | model.noteLike | witDetail | wit | rdgGrp )* )
}
```

---

**<appInfo>** (application information) records information about an application which has edited the TEI file.
[*2.3.11. The Application Information Element*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*

    * *@resp*
  – att.global.source
    * *@source*

*Member of* model.encodingDescPart
*Contained by*
*header:* encodingDesc
*May contain*
*header:* application
*Example*

```
<appInfo>
  <application version="1.24" ident="Xaira">
    <label>XAIRA Indexer</label>
    <ptr target="#P1"/>
  </application>
</appInfo>
```

*Content model*

```
<content>
 <classRef key="model.applicationLike"
  minOccurs="1" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element appInfo
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   model.applicationLike+
}
```

---

**<application>** provides information about an application which has acted upon the document. [*2.3.11. The Application Information Element*]

*Module* header — The TEI Header
*Attributes* • att.global
  – *@xml:id*
  – *@n*
  – *@xml:lang*
  – *@xml:base*
  – *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*

  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*

- att.typed
  - *@type*
  - *@subtype*

`@ident` supplies an identifier for the application, independent of its version number or display name.
*Status* Required
*Datatype* teidata.name

`@version` supplies a version number for the application, independent of its identifier or display name.
*Status* Required
*Datatype* teidata.versionNumber

*Member of* model.applicationLike
*Contained by*
*header:* appInfo
*May contain*
*core:* desc label p ptr ref
*linking:* ab
*tagdocs:* listRef
*Example*

```
<appInfo>
  <application  version="1.5"
    ident="ImageMarkupTool1"  notAfter="2006-06-01">
    <label>Image Markup Tool</label>
    <ptr  target="#P1"/>
    <ptr  target="#P2"/>
  </application>
</appInfo>
```

This example shows an appInfo element documenting the fact that version 1.5 of the Image Markup Tool1 application has an interest in two parts of a document which was last saved on June 6 2006. The parts concerned are accessible at the URLs given as target for the two <ptr> elements.

*Content model*

```
<content>
 <sequence>
  <classRef key="model.labelLike"
   minOccurs="1" maxOccurs="unbounded"/>
  <alternate>
   <classRef key="model.ptrLike"
    minOccurs="0" maxOccurs="unbounded"/>
   <classRef key="model.pLike"
    minOccurs="0" maxOccurs="unbounded"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element application
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.typed.attributes,
   attribute ident { text },
   attribute version { text },
   ( model.labelLike+, ( model.ptrLike* | model.pLike* ) )
}
```

---

**<arc>** (arc) encodes an arc, the connection from one node to another in a graph. [*20.1. Graphs and Digraphs*]

*Module* nets — Graphs, Networks, and Trees

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@from` gives the identifier of the node which is adjacent from this arc.
*Status* Required
*Datatype* teidata.pointer

`@to` gives the identifier of the node which is adjacent to this arc.
*Status* Required
*Datatype* teidata.pointer

*Contained by* graph

*May contain*

*core:* label

*Note* The <arc> element must be used if the arcs are labeled. Otherwise, arcs can be encoded using the *adj*, *adjTo* and *adjFrom* attributes on the <node> tags in the graph. Both <arc> tags and adjacency attributes can be used, but the resulting encoding would be highly redundant.

Zero, one, or two children <label> elements may be present. The first occurrence of <label> provides a label for the arc; the second provides a second label for the arc, and should be used if a transducer is being encoded.

*Example*

```
<arc  from="#T3"  to="#T3">
  <label>OLD</label>
  <label>VIEUX</label>
</arc>
```

*Content model*

```
<content>
 <sequence minOccurs="0">
  <elementRef key="label"/>
  <elementRef key="label" minOccurs="0"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element arc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute from { text },
```

```
   attribute to { text },
   ( ( label, label? )? )
}
```

---

**<argument>** (argument) contains a formal list or prose description of the topics addressed by a subdivision of a text. [*4.2. Elements Common to All Divisions 4.6. Title Pages*]

*Module* textstructure — Default Text Structure

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*

*Member of* model.divWrapper model.pLike.front model.titlepagePart

*Contained by*

*cmc:* post
*core:* lg list
*drama:* castList epilogue performance prologue
*figures:* figure table
*msdescription:* msItem
*textcrit:* lem rdg
*textstructure:* back body div div1 div2 div3 div4 div5 div6 div7 front group opener titlePage
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*cmc:* post
*core:* bibl biblStruct cb cit desc ellipsis gap gb head index l label lb lg list listBibl milestone note noteGrp p pb q quote said sp stage
*dictionaries:* entry entryFree superEntry
*drama:* camera caption castList move sound spGrp tech view
*figures:* figure notatedMusic table
*header:* biblFull
*iso-fs:* fLib fs fvLib
*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* msDesc
*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation
*nets:* eTree forest graph listForest tree
*spoken:* annotationBlock incident kinesic pause shift u vocal writing
*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef
*textcrit:* app listApp listWit witDetail
*textstructure:* floatingText
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
*Example*

```
<argument>
  <p>Monte Video — Maldonado — Excursion
     to R Polanco — Lazo and Bolas — Partridges —
     Absence of Trees — Deer — Capybara, or River Hog —
     Tucutuco — Molothrus, cuckoo-like habits — Tyrant
     Flycatcher — Mocking-bird — Carrion Hawks —
     Tubes formed by Lightning — House struck</p>
</argument>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.global"/>
   <classRef key="model.headLike"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.common"/>
   <classRef key="model.global"
```

```
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element argument
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   ( ( model.global | model.headLike )*, ( ( model.common, model.global* )+ ) )
}
```

---

**<att>** (attribute) contains the name of an attribute appearing within running text. [*23. Documentation Elements*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@scheme` (scheme) supplies an identifier for the scheme in which this name is defined.

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **TEI** (Text Encoding Initiative) this attribute is part of the TEI scheme.[Default]

**DBK** (Docbook) this attribute is part of the Docbook scheme.

**XX** (unknown) this attribute is part of an unknown scheme.

**imaginary** (imaginary) the attribute is from a non-existent scheme, for illustrative purposes only

**XHTML** (XHTML) the attribute is part of the XHTML language

**XML** (XML) the attribute is part of the XML language

**XI** (XI) the attribute is defined in the xInclude schema

*Member of* model.phrase.xml

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain* XSD Name

*Note* As an alternative to using the *scheme* attribute a namespace prefix may be used. Where both *scheme* and a prefix are used, the prefix takes precedence.

*Example*

```
<p>The TEI defines several <soCalled>global</soCalled> attributes; their names
include
<att>xml:id</att>, <att>rend</att>, <att>xml:lang</att>, <att>n</att>,
<att>xml:space</att>,
and <att>xml:base</att>; <att  scheme="XX">type</att> is not amongst them.</p>
```

*Content model* `<content> <dataRef key="teidata.name"/></content>`

*Schema Declaration*

```
element att
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute scheme { text }?,
   teidata.name}
```

---

## <attDef>

(attribute definition) contains the definition of a single attribute. [*23.5.3. Attribute List Specification*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

• att.identified
– *@ident*
– *@predeclare*
– *@module*

– att.combinable
* *@mode*

* att.deprecated
· *@validUntil*

• att.namespaceable
– *@ns*

`@usage` specifies the optionality of the attribute.
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **req** (required)
**rec** (recommended)
**opt** (optional) [Default]

*Contained by tagdocs:* attList
*May contain*
*core:* desc gloss
*tagdocs:* altIdent constraintSpec datatype defaultVal equiv exemplum remarks valDesc valList
*Example*

```
<attDef usage="rec" ident="type">
  <desc>specifies a name conventionally used for this level of subdivision,
e.g.
```

```
  <val>act</val>, <val>volume</val>, <val>book</val>, <val>section</val>,
<val>canto</val>,
    etc.</desc>
  <datatype>
    <dataRef  key="teidata.word"/>
  </datatype>
</attDef>
```

*Schematron* <sch:ns prefix="teix"  uri="http://www.tei-c.org/ns/Examples"/> <sch:rule context="tei:attDef"> <sch:assert test="ancestor::teix:egXML[ @valid eq 'feasible'] or @mode eq 'change' or @mode eq 'delete' or tei:datatype or tei:valList[ @type eq 'closed']"> Attribute: the definition of the @<sch:value-of select="@ident"/> attribute in the <sch:value-of select="ancestor::*[@ident][1]/@ident"/>  <sch:value-of select="' '"/>  <sch:value-of select="local-name(ancestor::*[@ident][1])"/> should have a closed valList or a datatype </sch:assert> </sch:rule>

*Schematron* <sch:rule context="tei:attDef[@usage eq 'req']"> <sch:report test="tei:defaultVal">Since the @<sch:value-of select="@ident"/> attribute is required, it will always be specified. Thus the default value (of "<sch:value-of select="normalize-space(tei:defaultVal)"/>") will never be used. Either change the definition of the attribute so it is not required ("rec" or "opt"), or remove the defaultVal element.</sch:report> </sch:rule>

*Schematron* <sch:rule context="tei:attDef[ tei:defaultVal and tei:valList[ @type eq 'closed'] and tei:datatype[ @maxOccurs > 1 or @minOccurs > 1 or @maxOccurs eq 'unbounded'] ]"> <sch:assert test="tokenize(normalize-space(tei:defaultVal),' ') = tei:valList/tei:valItem/@ident">In the <sch:value-of select="local-name(ancestor::*[@ident][1])"/> defining <sch:value-of select="ancestor::*[@ident][1]/@ident"/> the default value of the @<sch:value-of select="@ident"/> attribute is not among the closed list of possible values</sch:assert> </sch:rule>

*Schematron* <sch:rule context="tei:attDef[ tei:defaultVal and tei:valList[ @type eq 'closed'] and tei:datatype[ not(@maxOccurs) or ( if ( @maxOccurs castable as xs:integer ) then ( @maxOccurs cast as xs:integer eq 1 ) else false() ) ] ]"> <sch:assert test="string(tei:defaultVal) = tei:valList/tei:valItem/@ident">In the <sch:value-of select="local-name(ancestor::*[@ident][1])"/> defining <sch:value-of select="ancestor::*[@ident][1]/@ident"/> the default value of the @<sch:value-of select="@ident"/> attribute is not among the closed list of possible values</sch:assert> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identSynonyms"/>
   <classRef key="model.descLike"/>
  </alternate>
  <elementRef key="datatype" minOccurs="0"/>
  <elementRef key="constraintSpec"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="defaultVal"
   minOccurs="0"/>
  <alternate minOccurs="0">
   <elementRef key="valList"/>
   <elementRef key="valDesc" minOccurs="1"
    maxOccurs="unbounded"/>
  </alternate>
  <elementRef key="exemplum" minOccurs="0"
   maxOccurs="unbounded"/>
```

```
  <elementRef key="remarks" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element attDef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.identified.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   att.namespaceable.attributes,
   attribute usage { "req" | "rec" | "opt" }?,
   (
      ( model.identSynonyms | model.descLike )*,
      datatype?,
      constraintSpec*,
      defaultVal?,
      ( valList | valDesc+ )?,
      exemplum*,
      remarks*
   )
}
```

---

**<attList>** (attribute list) contains documentation for all the attributes associated with this element, as a series of <attDef> elements. [*23.5. Element Specifications 23.6. Class Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@org` (organization) specifies whether only one (choice) or all (group) of the attributes in the list are available.
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **group** grouped[Default]
**choice** alternated

*Contained by tagdocs:* attList classSpec elementSpec
*May contain*
*tagdocs:* attDef attList attRef
*Example*

```
<attList>
  <attDef ident="type" usage="opt">
    <desc>type of schema</desc>
    <datatype>
      <dataRef key="teidata.enumerated"/>
    </datatype>
  </attDef>
</attList>
```

*Example*

```
<attList org="choice">
  <attDef ident="active">
    <desc versionDate="2005-07-24"
      xml:lang="en">identifies the <soCalled>active</soCalled> participants in a
non-mutual relationship, or all the participants in a mutual one.</desc>
    <datatype maxOccurs="unbounded">
      <dataRef key="teidata.pointer"/>
    </datatype>
```

```
      </attDef>
      <attDef ident="mutual" usage="opt">
        <desc versionDate="2005-07-24"
          xml:lang="en">supplies a list of participants amongst all of whom the
    relationship holds equally.</desc>
        <datatype maxOccurs="unbounded">
          <dataRef key="teidata.pointer"/>
        </datatype>
      </attDef>
    </attList>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <elementRef key="attRef"/>
  <elementRef key="attDef"/>
  <elementRef key="attList"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element attList
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute org { "group" | "choice" }?,
   ( attRef | attDef | attList )+
}
```

---

**<attRef>** (attribute pointer) points to the definition of an attribute or group of attributes. [*23.2. Modules and Schemas*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@class` the name of the attribute class
*Status* Optional
*Datatype* teidata.name

`@name` the name of the attribute
*Status* Optional
*Datatype* teidata.name

*Contained by:* attLists: attList
*May contain* Empty element
*Example*

```
<attRef class="att.global" name="xml:id"/>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element attRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
```

```
   attribute class { text }?,
   attribute name { text }?,
   empty
}
```

---

**<author>** (author) in a bibliographic reference, contains the name(s) of an author, personal or corporate, of a work; for example in the same form as that provided by a recognized bibliographic name authority. [*3.12.2.2. Titles, Authors, and Editors 2.2.1. The Title Statement*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
    * *@key*
    * *@ref*

*Member of* model.respLike

*Contained by*

*core:* analytic bibl monogr

*header:* editionStmt titleStmt

*msdescription:* msItem msItemStruct

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Particularly where cataloguing is likely to be based on the content of the header, it is advisable to use a generally recognized name authority file to supply the content for this element. The attributes *key* or *ref* may also be used to reference canonical information about the author(s) intended from any appropriate authority, such as a library catalogue or online resource.

In the case of a broadcast, use this element for the name of the company or network responsible for making the broadcast.

Where an author is unknown or unspecified, this element may contain text such as *Unknown* or *Anonymous*. When the appropriate TEI modules are in use, it may also contain detailed tagging of the names used for people, organizations or places, in particular where multiple names are given.

*Example*

```
<author>British Broadcasting Corporation</author>
<author>La Fayette, Marie Madeleine Pioche de la Vergne, comtesse de
(1634–1693)</author>
<author>Anonymous</author>
<author>Bill and Melinda Gates Foundation</author>
<author>
  <persName>Beaumont, Francis</persName>
and
<persName>John Fletcher</persName>
</author>
<author>
  <orgName  key="BBC">British Broadcasting
     Corporation</orgName>: Radio 3 Network
</author>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element author
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
```

```
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   macro.phraseSeq}
```

---

**<authority>** (release authority) supplies the name of a person or other agency responsible for making a work available, other than a publisher or distributor. [*2.2.4. Publication, Distribution, Licensing, etc.*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

    * *@source*

  - att.canonical
    - *@key*
    - *@ref*

*Member of* model.publicationStmtPart.agency

*Contained by*

*core:* monogr

*header:* publicationStmt

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin

character data

*Example*

```
<authority>John Smith</authority>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element authority
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
```

```
   att.canonical.attributes,
   macro.phraseSeq.limited}
```

---

**<availability>** (availability) supplies information about the availability of a text, for example any restrictions on its use or distribution, its copyright status, any licence applying to it, etc. [*2.2.4. Publication, Distribution, Licensing, etc.*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable
  - *@default*

`@status` (status) supplies a code identifying the current availability of the text.

*Status* Optional

*Datatype* teidata.enumerated
*Legal values are:* **free** (free) the text is freely available.
**unknown** (unknown) the status of the text is unknown.
**restricted** (restricted) the text is not freely available.

*Member of* model.biblPart model.publicationStmtPart.detail
*Contained by*
*core:* analytic bibl monogr series
*header:* publicationStmt
*msdescription:* adminInfo
*May contain*
*core:* p
*header:* licence
*linking:* ab
*Note* A consistent format should be adopted
*Example*

```
<availability status="restricted">
  <p>Available for academic research purposes only.</p>
</availability>
<availability status="free">
  <p>In the public domain</p>
</availability>
<availability status="restricted">
  <p>Available under licence from the publishers.</p>
</availability>
```

*Example*

```
<availability>
  <licence target="http://opensource.org/licenses/MIT">
    <p>The MIT License
        applies to this document.</p>
    <p>Copyright (C) 2011 by The University of Victoria</p>
    <p>Permission is hereby granted, free of charge, to any person obtaining a
copy
        of this software and associated documentation files (the "Software"),
to deal
        in the Software without restriction, including without limitation the
rights
        to use, copy, modify, merge, publish, distribute, sublicense, and/or
sell
        copies of the Software, and to permit persons to whom the Software is
        furnished to do so, subject to the following conditions:</p>
    <p>The above copyright notice and this permission notice shall be included
in
        all copies or substantial portions of the Software.</p>
    <p>THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS
OR
        IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF
MERCHANTABILITY,
        FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL
THE
        AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER
        LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING
```

```
FROM,
       OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS
IN
       THE SOFTWARE.</p>
  </licence>
</availability>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:availability"/> </sch:pattern>
*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.availabilityPart"/>
  <classRef key="model.pLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element availability
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   attribute status { "free" | "unknown" | "restricted" }?,
   ( model.availabilityPart | model.pLike )+
}
```

---

## <back>

<back> (back matter) contains any appendixes, etc. following the main part of a text. [*4.7. Back Matter 4. Default Text Structure*]

*Module* textstructure — Default Text Structure
*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.declaring
  - *@decls*

*Contained by* *tei:textstructure:* floatingText text
*transcr:* facsimile
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* cb divGen ellipsis gap gb head index lb list listBibl milestone note noteGrp p pb
*drama:* castList epilogue performance prologue set
*figures:* figure notatedMusic table
*iso-fs:* fLib fs fvLib
*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline
*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* schemaSpec
*textcrit:* app listApp listWit witDetail
*textstructure:* argument byline closer dateline div div1 docAuthor docDate docEdition docImprint docTitle epigraph postscript signed titlePage titlePart trailer
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
*Note* Because cultural conventions differ as to which elements are grouped as back matter and which as front matter, the content models for the <back> and <front> elements are identical.
*Example*

```
<back>
  <div  type="appendix">
    <head>The Golden Dream or, the Ingenuous Confession</head>
    <p>TO shew the Depravity of human Nature, and how apt the Mind is to be
misled by Trinkets
        and false Appearances, Mrs. Two-Shoes does acknowledge, that after she
became rich, she
        had like to have been, too fond of Money
<!-- .... -->
    </p>
  </div>
<!-- ... -->
  <div  type="epistle">
    <head>A letter from the Printer, which he desires may be inserted</head>
    <salute>Sir.</salute>
    <p>I have done with your Copy, so you may return it to the Vatican, if you
please;

<!-- ... -->
    </p>
  </div>
  <div  type="advert">
    <head>The Books usually read by the Scholars of Mrs Two-Shoes are these and
are sold at Mr
        Newbery's at the Bible and Sun in St Paul's Church-yard.</head>
    <list>
      <item  n="1">The Christmas Box, Price 1d.</item>
      <item  n="2">The History of Giles Gingerbread, 1d.</item>
<!-- ... -->
      <item  n="42">A Curious Collection of Travels, selected from the Writers
of all Nations,
            10 Vol, Pr. bound 1l.</item>
    </list>
  </div>
  <div  type="advert">
    <head>By the KING's Royal Patent, Are sold by J. NEWBERY, at the Bible and
Sun in St.
        Paul's Church-Yard.</head>
    <list>
      <item  n="1">Dr. James's Powders for Fevers, the Small-Pox, Measles,
Colds, &c. 2s.
            6d</item>
      <item  n="2">Dr. Hooper's Female Pills, 1s.</item>
<!-- ... -->
    </list>
  </div>
</back>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.frontPart"/>
   <classRef key="model.pLike.front"/>
   <classRef key="model.pLike"/>
   <classRef key="model.listLike"/>
```

```
   <classRef key="model.global"/>
  </alternate>
  <alternate minOccurs="0">
   <sequence>
    <classRef key="model.div1Like"/>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <classRef key="model.frontPart"/>
     <classRef key="model.div1Like"/>
     <classRef key="model.global"/>
    </alternate>
   </sequence>
   <sequence>
    <classRef key="model.divLike"/>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <classRef key="model.frontPart"/>
     <classRef key="model.divLike"/>
     <classRef key="model.global"/>
    </alternate>
   </sequence>
  </alternate>
  <sequence minOccurs="0">
   <classRef key="model.divBottomPart"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.divBottomPart"/>
    <classRef key="model.global"/>
   </alternate>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element back
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   (
      (
         model.frontPart       | model.pLike.front       | model.pLike       | model.listL
      (
         (
            model.div1Like,
            ( model.frontPart | model.div1Like | model.global )*
         )
       | ( model.divLike, ( model.frontPart | model.divLike | model.global )* )
      )?,
      ( ( model.divBottomPart, ( model.divBottomPart | model.global )* )? )
```

```
    )
}
```

---

**<bibl>** (bibliographic citation) contains a loosely-structured bibliographic citation of which the sub-components may or may not be explicitly tagged. [*3.12.1. Methods of Encoding Bibliographic References and Lists of References 2.2.7. The Source Description 16.3.2. Declarable Elements*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.cmc

  - *@generatedBy*
- att.declarable
  - *@default*
- att.docStatus
  - *@status*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.biblLike model.biblPart

*Contained by*

*cmc:* post

*core:* add bibl cit corr del desc emph head hi item l listBibl meeting note orig p q quote ref reg relatedItem said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote sourceDesc tagUsage taxonomy

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout msItem msItemStruct musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* climate event location object occupation org person personGrp persona place population state terrain trait

*spoken:* broadcast scriptStmt u writing

*textcrit:* lem rdg witDetail witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address author bibl biblScope cb choice citedRange corr date del distinct editor ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp meeting mentioned milestone name note noteGrp num orig pb ptr pubPlace publisher q quote ref reg relatedItem respStmt rs series sic soCalled term textLang time title unclear unit

*dictionaries:* lang

*figures:* figure notatedMusic

*gaiji:* g

*header:* availability distributor edition extent funder idno principal sponsor

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* decoNote depth dim explicit height incipit msIdentifier width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listRelation location nameLink objectIdentifier objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* code ident listRef

*textcrit:* app witDetail

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

character data

*Note* Contains *phrase-level* elements, together with any combination of elements from the `model.biblPart` class

*Example*

```
<bibl>Blain, Clements and Grundy: Feminist Companion to Literature in English
(Yale,
1990)</bibl>
```

*Example*

```
<bibl>
  <title  level="a">The Interesting story of the Children in the Wood</title>.
In
<author>Victor E Neuberg</author>, <title>The Penny Histories</title>.
<publisher>OUP</publisher>
  <date>1968</date>.
</bibl>
```

*Example*

```
<bibl  type="article"  subtype="book_chapter"
  xml:id="carlin_2003">
  <author>
    <name>
      <surname>Carlin</surname>
         (<forename>Claire</forename>)</name>
  </author>,
<title  level="a">The Staging of Impotence : France’s last
    congrès</title> dans
<bibl  type="monogr">
   <title  level="m">Theatrum mundi : studies in honor of Ronald W.
       Tobin</title>, éd.
  <editor>
     <name>
       <forename>Claire</forename>
       <surname>Carlin</surname>
     </name>
  </editor> et
  <editor>
     <name>
       <forename>Kathleen</forename>
       <surname>Wine</surname>
     </name>
  </editor>,
  <pubPlace>Charlottesville, Va.</pubPlace>,
  <publisher>Rookwood Press</publisher>,
  <date  when="2003">2003</date>.
```

```
      </bibl>
    </bibl>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:bibl"/> </sch:pattern>

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.highlighted"/>
  <classRef key="model.pPart.data"/>
  <classRef key="model.pPart.edit"/>
  <classRef key="model.segLike"/>
  <classRef key="model.ptrLike"/>
  <classRef key="model.biblPart"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element bibl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.cmc.attributes,
   att.declarable.attributes,
   att.docStatus.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      text
    | model.gLike    | model.highlighted    | model.pPart.data    | model.pPart.edit    |
}
```

---

**<biblFull>** (fully-structured bibliographic citation) contains a fully-structured bibliographic citation, in which all components of the TEI file description are present. [*3.12.1. Methods of Encoding Bibliographic References and Lists of References 2.2. The File Description 2.2.7. The Source Description 16.3.2. Declarable Elements*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.canonical
  - *@key*
  - *@ref*
- att.cmc
  - *@generatedBy*
- att.declarable
  - *@default*
- att.docStatus
  - *@status*
- att.sortable
  - *@sortKey*

*Member of* model.biblLike

*Contained by*

*cmc:* post

*core:* add cit corr del desc emph head hi item l listBibl meeting note orig p q quote ref reg relatedItem said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote sourceDesc tagUsage taxonomy

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout msItem musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* climate event location object occupation org person personGrp persona place population state terrain trait

*spoken:* broadcast scriptStmt u writing

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*header:* editionStmt extent fileDesc notesStmt profileDesc publicationStmt seriesStmt sourceDesc titleStmt

*Example*

```
<biblFull>
  <titleStmt>
    <title>The Feminist Companion to Literature in English: women writers from
the middle ages
        to the present</title>
    <author>Blain, Virginia</author>
    <author>Clements, Patricia</author>
    <author>Grundy, Isobel</author>
  </titleStmt>
  <editionStmt>
    <edition>UK edition</edition>
  </editionStmt>
  <extent>1231 pp</extent>
  <publicationStmt>
    <publisher>Yale University Press</publisher>
    <pubPlace>New Haven and London</pubPlace>
    <date>1990</date>
  </publicationStmt>
  <sourceDesc>
    <p>No source: this is an original work</p>
  </sourceDesc>
</biblFull>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:biblFull"/> </sch:pattern>

*Content model*

```
<content>
 <alternate>
  <sequence>
   <sequence>
    <elementRef key="titleStmt"/>
    <elementRef key="editionStmt"
     minOccurs="0"/>
    <elementRef key="extent" minOccurs="0"/>
    <elementRef key="publicationStmt"/>
    <elementRef key="seriesStmt"
```

```
     minOccurs="0" maxOccurs="unbounded"/>
    <elementRef key="notesStmt"
     minOccurs="0"/>
   </sequence>
   <elementRef key="sourceDesc"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <sequence>
   <elementRef key="fileDesc"/>
   <elementRef key="profileDesc"/>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element biblFull
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.cmc.attributes,
   att.declarable.attributes,
   att.docStatus.attributes,
   att.sortable.attributes,
   (
      (
         (
            titleStmt,
            editionStmt?,
            extent?,
            publicationStmt,
            seriesStmt*,
            notesStmt?
         ),
         sourceDesc*
      )
    | ( fileDesc, profileDesc )
   )
}
```

---

## <biblScope>

(scope of bibliographic reference) defines the scope of a bibliographic reference, for example as a list of page numbers, or a named subdivision of a larger work. [*3.12.2.5. Scopes and Ranges in Bibliographic Citations*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.citing
  - *@unit*
  - *@from*
  - *@to*

*Member of* model.imprintPart

*Contained by*

*core:* bibl imprint monogr series

*header:* seriesStmt

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* When a single page is being cited, use the *from* and *to* attributes with an identical value. When no clear endpoint is provided, the *from* attribute may be used without *to*; for example a citation such as 'p. 3ff' might be encoded `<biblScope from="3">p. 3ff</biblScope>`.

It is now considered good practice to supply this element as a sibling (rather than a child) of <imprint>, since it supplies information which does not constitute part of the imprint.

*Example*

```
<biblScope>pp 12–34</biblScope>
<biblScope  unit="page"  from="12"  to="34"/>
<biblScope  unit="volume">II</biblScope>
<biblScope  unit="page">12</biblScope>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element biblScope
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.citing.attributes,
   macro.phraseSeq}
```

**<biblStruct>** (structured bibliographic citation) contains a structured bibliographic citation, in which only bibliographic sub-elements appear and in a specified order. [*3.12.1. Methods of Encoding Bibliographic References and Lists of References 2.2.7. The Source Description 16.3.2. Declarable Elements*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.cmc
  - *@generatedBy*
- att.declarable
  - *@default*
- att.docStatus
  - *@status*

- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.biblLike

*Contained by*

*cmc:* post

*core:* add cit corr del desc emph head hi item l listBibl meeting note orig p q quote ref reg relatedItem said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote sourceDesc tagUsage taxonomy

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout msItem msItemStruct musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* climate event location object occupation org person personGrp persona place population state terrain trait

*spoken:* broadcast scriptStmt u writing

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* analytic citedRange monogr note noteGrp ptr ref relatedItem series

*tagdocs:* listRef

*Example*

```
<biblStruct>
  <monogr>
    <author>Blain, Virginia</author>
    <author>Clements, Patricia</author>
    <author>Grundy, Isobel</author>
    <title>The Feminist Companion to Literature in English: women writers from
the middle ages
        to the present</title>
    <edition>first edition</edition>
    <imprint>
      <publisher>Yale University Press</publisher>
      <pubPlace>New Haven and London</pubPlace>
      <date>1990</date>
    </imprint>
  </monogr>
</biblStruct>
```

*Example*

```
<biblStruct  type="newspaper">
  <analytic>
    <author>
      <forename>David</forename>
      <surname>Barstow</surname>
    </author>
    <author>
      <forename>Susanne</forename>
      <surname>Craig</surname>
    </author>
    <author>
      <forename>Russ</forename>
      <surname>Buettner</surname>
    </author>
    <title  type="main">Trump Took Part in Suspect Schemes to Evade Tax
Bills</title>
    <title  type="sub">Behind the Myth of a Self-Made Billionaire, a Vast
Inheritance From His Father</title>
  </analytic>
  <monogr>
    <title  level="j">The New York Times</title>
    <imprint>
      <pubPlace>New York</pubPlace>
      <publisher>A. G. Sulzberger</publisher>
      <date  when="2018-10-03">Wednesday, October 3, 2018</date>
    </imprint>
    <biblScope  unit="volume">CLXVIII</biblScope>
    <biblScope  unit="issue">58,104</biblScope>
    <biblScope  unit="page">1</biblScope>
  </monogr>
</biblStruct>
```

*Schematron* <sch:pattern is-a=”declarable”> <sch:param name=”tde” value=”tei:biblStruct”/> </sch:pattern>

*Content model*

```
<content>
 <sequence>
  <elementRef key="analytic" minOccurs="0"
   maxOccurs="unbounded"/>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="monogr"/>
   <elementRef key="series" minOccurs="0"
    maxOccurs="unbounded"/>
  </sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.noteLike"/>
   <classRef key="model.ptrLike"/>
   <elementRef key="relatedItem"/>
   <elementRef key="citedRange"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element biblStruct
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.cmc.attributes,
   att.declarable.attributes,
   att.docStatus.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      analytic*,
      ( ( monogr, series* )+ ),
      ( model.noteLike | model.ptrLike | relatedItem | citedRange )*
   )
}
```

---

**<bicond>** (bi-conditional feature-structure constraint) defines a biconditional feature-structure constraint; both consequent and antecedent are specified as feature structures or groups of feature structures; the constraint is satisfied if both subsume a given feature structure, or if both do not. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Contained by:* *iso-fs:* fsConstraints
*May contain*
*iso-fs:* f fs iff
*Example*

```
<bicond>
  <fs>
    <f  name="FOO">
      <symbol  value="42"/>
    </f>
  </fs>
  <iff/>
  <fs>
    <f  name="BAR">
      <binary  value="true"/>
    </f>
  </fs>
</bicond>
```

*Content model*

```
<content>
 <sequence>
  <alternate>
   <elementRef key="fs"/>
   <elementRef key="f"/>
  </alternate>
  <elementRef key="iff"/>
  <alternate>
   <elementRef key="fs"/>
   <elementRef key="f"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element bicond
{
```

```
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( ( fs | f ), iff, ( fs | f ) )
}
```

---

**<binary>** (binary value) represents the value part of a feature-value specification which can contain either of exactly two possible values. [*19.2. Elementary Feature Structures and the Binary Feature Value*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

    * *@source*

  - att.datcat
    - *@datcat*
    - *@valueDatcat*
    - *@targetDatcat*

  `@value` supplies a binary value.
    *Status* Required
    *Datatype* teidata.truthValue
    *Note* This attribute has a datatype of `teidata.truthValue`, which may be represented by the values 1 or true, or 0 or false.

*Member of* model.featureVal.single

*Contained by*

*iso-fs:* f fvLib if vAlt vColl vDefault vLabel vMerge vNot vRange

*May contain* Empty element

*Note* The value attribute may take any value permitted for attributes of the W3C datatype Boolean: this includes for example the strings `true` or `1` which are equivalent.

*Example*

```
<f name="strident">
  <binary value="true"/>
</f>
<f name="exclusive">
  <binary value="false"/>
</f>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element binary
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datcat.attributes,
   attribute value { text },
   empty
}
```

---

## <binaryObject> provides encoded binary data representing an inline graphic, audio, video or other object.

[*3.10. Graphics and Other Non-textual Components*]

*Module* core — Elements Available in All TEI Documents

*Attributes* • att.global
  - *@xml:id*
  - *@n*

  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.media
  - *@width*
  - *@height*
  - *@scale*

  - att.internetMedia
    * *@mimeType*
- att.timed
  - *@start*
  - *@end*

  - att.duration
    * att.duration.iso

- *@dur-iso*

* att.duration.w3c
- *@dur*

- att.typed
  - *@type*
  - *@subtype*

`@encoding` The encoding used to encode the binary data. If not specified, this is assumed to be Base64.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

*Member of* model.graphicLike model.titlepagePart
*Contained by*
*analysis:* cl phr s
*cmc:* post
*core:* abbr add addrLine author biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castItem role roleDesc sound tech view
*figures:* cell figure formula notatedMusic table
*gaiji:* char glyph
*header:* change distributor edition extent geoDecl handNote licence scriptNote
*linking:* ab seg
*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material msItem musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark
*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname
*spoken:* u writing
*tagdocs:* eg
*textcrit:* lem rdg wit witDetail
*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePage titlePart trailer
*transcr:* damage facsimile fw metamark mod restore retrace secl sourceDoc supplied surface surplus zone
*verse:* rhyme
*May contain* Character data only
*Example*

```
<binaryObject mimeType="image/gif">
R0lGODdhMAAwAPAAAAAAAP///ywAAAAAMAAwAAAC8IyPqcvt3wCcDkiLc7C0qwy
GHhSWpjQu5yqmCYsapyuvUUlvONmOZtfzgFzByTB10QgxOR0TqBQejhRNzOfkVJ
+5YiUqrXF5Y5lKh/DeuNcP5yLWGsEbtLiOSpa/TPg7JpJHxyendzWTBfX0cxOnK
PjgBzi4diinWGdkF8kjdfnycQZXZeYGejmJlZeGl9i2icVqaNVailT6F5iJ90m6
mvuTS4OK05M0vDk0Q4XUtwvKOzrcd3iq9uisF81M1OIcR7lEewwcLp7tuNNkM3u
Nna3F2JQFo97Vriy/Xl4/f1cf5VWzXyym7PH hhx4dbgYKAAA7</binaryObject>
```

*Content model* `<content> <textNode/></content>`

*Schema Declaration*

```
element binaryObject
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.media.attributes,
   att.internetMedia.attributes,
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   att.typed.attributes,
   attribute encoding { list { + } }?,
   text
}
```

---

## <binding>

<binding> (binding) contains a description of one binding, i.e. type of covering, boards, etc. applied to a manuscript or other object. [*11.7.3.1. Binding Descriptions*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

  * *@select*
  - att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
  - att.global.responsibility
  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.datable
  - *@period*
  - att.datable.custom
  * *@when-custom*
  * *@notBefore-custom*
  * *@notAfter-custom*
  * *@from-custom*
  * *@to-custom*
  * *@datingPoint*
  * *@datingMethod*
  - att.datable.iso
  * *@when-iso*
  * *@notBefore-iso*
  * *@notAfter-iso*
  * *@from-iso*
  * *@to-iso*
  - att.datable.w3c
  * *@when*
  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*

`@contemporary` (contemporary) specifies whether or not the binding is contemporary with the majority of its contents.

*Status* Optional

*Datatype* teidata.xTruthValue

*Note* The value true indicates that the binding is contemporaneous with its contents; the value false that it is not. The value unknown should be used when the date of either binding or manuscript is unknown

*Contained by:* bindingDesc

*May contain*

*core:* p

*linking:* ab
*msdescription:* condition decoNote
*Example*

```
<binding  contemporary="true">
  <p>Contemporary blind stamped leather over wooden boards with evidence of a
fore edge clasp
    closing to the back cover.</p>
</binding>
```

*Example*

```
<bindingDesc>
  <binding  contemporary="false">
    <p>Quarter bound by the Phillipps' binder, Bretherton, with his sticker on
the front
       pastedown.</p>
  </binding>
  <binding  contemporary="false">
    <p>Rebound by an unknown 19th c. company; edges cropped and gilt.</p>
  </binding>
</bindingDesc>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.pLike"/>
  <elementRef key="condition"/>
  <elementRef key="decoNote"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element binding
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   attribute contemporary { text }?,
   ( model.pLike | condition | decoNote )+
}
```

**<bindingDesc>** (binding description) describes the present and former bindings of a manuscript or other object, either as a series of paragraphs or as a series of distinct <binding> elements, one for each binding of the manuscript. [*11.7.3.1. Binding Descriptions*]

*Module* msdescription — Manuscript Description

*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  - * *@ana*
- – att.global.change
  - * *@change*
- – att.global.facs
  - * *@facs*
- – att.global.linking
  - * *@corresp*
  - * *@synch*
  - * *@sameAs*
  - * *@copyOf*
  - * *@next*
  - * *@prev*
  - * *@exclude*
  - * *@select*
- – att.global.rendition
  - * *@rend*
  - * *@style*
  - * *@rendition*
- – att.global.responsibility
  - * *@cert*
  - * *@resp*
- – att.global.source
  - * *@source*

*Member of* model.physDescPart

*Contained by*

*msdescription:* physDesc

*May contain*

*core:* p

*linking:* ab

*msdescription:* binding condition decoNote

*Example*

```
<bindingDesc>
  <p>Sewing not visible; tightly rebound over
     19th-cent. pasteboards, reusing panels of 16th-cent. brown leather with
     gilt tooling à la fanfare, Paris c. 1580-90, the centre of each
     cover inlaid with a 17th-cent. oval medallion of red morocco tooled in
     gilt (perhaps replacing the identifying mark of a previous owner); the
     spine similarly tooled, without raised bands or title-piece; coloured
     endbands; the edges of the leaves and boards gilt.Boxed.</p>
</bindingDesc>
```

*Content model*

```
<content>
 <alternate>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.pLike"/>
   <elementRef key="decoNote"/>
   <elementRef key="condition"/>
  </alternate>
  <elementRef key="binding" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element bindingDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( ( model.pLike | decoNote | condition )+ | binding+ )
}
```

---

## <birth>

<birth> (birth) contains information about a person's birth, such as its date and place. [*16.2.2. The Participant Description*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*

  * *@notAfter*
  * *@from*
  * *@to*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
  * *@atLeast*
  * *@atMost*
  * *@min*
  * *@max*
  * *@confidence*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*

  - att.canonical
  * *@key*
  * *@ref*
- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **caesarean** (caesarean section)
**vaginal** (vaginal delivery)
**exNihilo** (ex nihilo)
**incorporated**
**founded**
**established**

*Member of* model.personPart
*Contained by*
*namesdates:* person personGrp persona
*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<birth>Before 1920, Midlands region.</birth>
```

*Example*

```
<birth  when="1960-12-10">In a small cottage near
<name  type="place">Aix-la-Chapelle</name>,
early in the morning of <date>10 Dec 1960</date>
</birth>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element birth
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
att.global.source.attributes,
att.datable.attributes,
att.datable.custom.attributes,
att.datable.iso.attributes,
att.datable.w3c.attributes,
att.dimensions.attributes,
att.ranging.attributes,
att.editLike.attributes,
att.naming.attributes,
att.canonical.attributes,
att.typed.attribute.subtype,
attribute type { text }?,
macro.phraseSeq}
```

---

## <bloc>

(bloc) contains the name of a geo-political unit consisting of two or more nation states or countries. [*14.2.3. Place Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*

    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
    * *@key*
    * *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.placeNamePart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num

orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage unitDef

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions altIdentifier catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msIdentifier musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectIdentifier objectName occupation offset org orgName persName persPronouns place placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<bloc type="union">the European Union</bloc>
<bloc type="continent">Africa</bloc>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element bloc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

## <body>

(text body) contains the whole body of a single unitary text, excluding any front or back matter. [*4. Default Text Structure*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change

    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.declaring
  - *@decls*

*Contained by:* tei:text floatingText text
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*cmc:* post
*core:* bibl biblStruct cb cit desc divGen ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage
*dictionaries:* entry entryFree superEntry
*drama:* camera caption castList move sound spGrp tech view
*figures:* figure notatedMusic table
*header:* biblFull
*iso-fs:* fLib fs fvLib
*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* msDesc
*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation
*nets:* eTree forest graph listForest tree
*spoken:* annotationBlock incident kinesic pause shift u vocal writing
*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition schemaSpec specGrp specGrpRef
*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline div div1 docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Example*

```
<body>
  <l>Nu scylun hergan hefaenricaes uard</l>
  <l>metudæs maecti end his modgidanc</l>
  <l>uerc uuldurfadur sue he uundra gihuaes</l>
  <l>eci dryctin or astelidæ</l>
  <l>he aerist scop aelda barnum</l>
  <l>heben til hrofe haleg scepen.</l>
  <l>tha middungeard moncynnæs uard</l>
  <l>eci dryctin æfter tiadæ</l>
  <l>firum foldu frea allmectig</l>
  <trailer>primo cantauit Cædmon istud carmen.</trailer>
</body>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.global"
   minOccurs="0" maxOccurs="unbounded"/>
  <sequence minOccurs="0">
   <classRef key="model.divTop"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.global"/>
    <classRef key="model.divTop"/>
   </alternate>
  </sequence>
  <sequence minOccurs="0">
   <classRef key="model.divGenLike"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.global"/>
    <classRef key="model.divGenLike"/>
   </alternate>
  </sequence>
  <alternate>
   <sequence minOccurs="1"
    maxOccurs="unbounded">
    <classRef key="model.divLike"/>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <classRef key="model.global"/>
     <classRef key="model.divGenLike"/>
    </alternate>
   </sequence>
   <sequence minOccurs="1"
    maxOccurs="unbounded">
    <classRef key="model.div1Like"/>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <classRef key="model.global"/>
     <classRef key="model.divGenLike"/>
    </alternate>
```

```
   </sequence>
   <sequence>
    <sequence minOccurs="1"
     maxOccurs="unbounded">
     <alternate minOccurs="1" maxOccurs="1">
      <elementRef key="schemaSpec"/>
      <classRef key="model.common"/>
     </alternate>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
    <alternate minOccurs="0">
     <sequence minOccurs="1"
      maxOccurs="unbounded">
      <classRef key="model.divLike"/>
      <alternate minOccurs="0"
       maxOccurs="unbounded">
       <classRef key="model.global"/>
       <classRef key="model.divGenLike"/>
      </alternate>
     </sequence>
     <sequence minOccurs="1"
      maxOccurs="unbounded">
      <classRef key="model.div1Like"/>
      <alternate minOccurs="0"
       maxOccurs="unbounded">
       <classRef key="model.global"/>
       <classRef key="model.divGenLike"/>
      </alternate>
     </sequence>
    </alternate>
   </sequence>
  </alternate>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divBottom"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element body
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   (
      model.global*,
      ( ( model.divTop, ( model.global | model.divTop )* )? ),
```

```
      ( ( model.divGenLike, ( model.global | model.divGenLike )* )? ),
      (
         ( ( model.divLike, ( model.global | model.divGenLike )* )+ )
       | ( ( model.div1Like, ( model.global | model.divGenLike )* )+ )
       | (
            ( ( ( schemaSpec | model.common ), model.global* )+ ),
            (
               ( ( model.divLike, ( model.global | model.divGenLike )* )+ )
             | ( ( model.div1Like, ( model.global | model.divGenLike )* )+ )
            )?
         )
      ),
      ( ( model.divBottom, model.global* )* )
   )
}
```

---

**<broadcast>** (broadcast) describes a broadcast used as the source of a spoken text. [*8.2. Documenting the Source of Transcribed Speech 16.3.2. Declarable Elements*]

*Module* spoken — Transcriptions of Speech

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.declarable
  - *@default*

*Member of* model.recordingPart
*Contained by*
*spoken:* recording
*May contain*
*core:* bibl biblStruct listBibl p
*header:* biblFull
*linking:* ab
*msdescription:* msDesc
*spoken:* recording
*Example*

```
<broadcast>
  <bibl>
    <author>Radio Trent</author>
    <title>Gone Tomorrow</title>
    <respStmt>
      <resp>Presenter</resp>
      <name>Tim Maby</name>
    </respStmt>
    <respStmt>
      <resp>Producer</resp>
      <name>Mary Kerr</name>
    </respStmt>
    <date  when="1989-06-12T12:30:00">12 June 89, 1230 pm</date>
  </bibl>
</broadcast>
```

*Schematron* <sch:pattern is-a=”declarable”> <sch:param name=”tde” value=”tei:broadcast”/> </sch:pattern>
*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <classRef key="model.biblLike"/>
  <elementRef key="recording"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element broadcast
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
```

```
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   ( model.pLike+ | model.biblLike | recording )
}
```

---

**<byline>** (byline) contains the primary statement of responsibility given for a work on its title page or at the head or end of the work. [*4.2.2. Openers and Closers 4.5. Front Matter*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc

– *@generatedBy*

*Member of* model.divWrapper model.pLike.front model.titlepagePart

*Contained by*

*cmc:* post

*core:* lg list

*drama:* castList epilogue performance prologue

*figures:* figure table

*msdescription:* msItem

*textcrit:* lem rdg

*textstructure:* back body closer div div1 div2 div3 div4 div5 div6 div7 front group opener titlePage

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q ref reg rs ruby sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* docAuthor

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The byline on a title page may include either the name or a description for the document's author. Where the name is included, it may optionally be tagged using the <docAuthor> element.

*Example*

```
<byline>Written by a CITIZEN who continued all the
while in London. Never made publick before.</byline>
```

*Example*

```
<byline>Written from her own MEMORANDUMS</byline>
```

*Example*

```
<byline>By George Jones, Political Editor, in Washington</byline>
```

*Example*

```
<byline>BY
<docAuthor>THOMAS PHILIPOTT,</docAuthor>
Master of Arts,
(Somtimes)
Of Clare-Hall in Cambridge.</byline>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <elementRef key="docAuthor"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element byline
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   ( text | model.gLike | model.phrase | docAuthor | model.global )*
}
```

---

**<c>** (character) represents a character. [*18.1. Linguistic Segment Categories*]

*Module* analysis — Simple Analytic Mechanisms

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change

    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.notated
  - *@notation*
- att.segLike
  - *@function*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.fragmentable
    * *@part*
  - att.metrical
    * *@met*
    * *@real*
    * *@rhyme*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.linePart model.segLike
*Contained by*

*analysis:* cl m pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain*

*gaiji:* g

character data

*Note* Contains a single character, a <g> element, or a sequence of graphemes to be treated as a single character. The *type* attribute is used to indicate the function of this segmentation, taking values such as letter, punctuation, or digit etc.

*Example*

```
<phr>
  <c>M</c>
  <c>O</c>
  <c>A</c>
  <c>I</c>
  <w>doth</w>
  <w>sway</w>
  <w>my</w>
  <w>life</w>
</phr>
```

*Content model* `<content> <macroRef key="macro.xtext"/></content>`

*Schema Declaration*

```
element c
{
   att.global.attributes,
   att.global.analytic.attributes,
```

```
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.notated.attributes,
   att.segLike.attributes,
   att.datcat.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   macro.xtext}
```

---

**<caesura>** marks the point at which a metrical line may be divided. [*6.2. Components of the Verse Line*]

*Module* verse — Verse

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*

- att.global.source
  - *@source*

*Member of* model.lPart

*Contained by*

*analysis:* cl phr s w

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain* Empty element

*Example*

```
<l>Hwæt we Gar-Dena <caesura/> in gear-dagum</l>
<l>þeod-cyninga <caesura/> þrym gefrunon,</l>
<l>hy ða æþelingas <caesura/> ellen fremedon.</l>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element caesura
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   empty
}
```

---

**<calendar>** (calendar) describes a calendar or dating system used in a dating formula in the text. [*2.4.5. Calendar Description*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*
- att.typed
  - *@type*
  - *@subtype*

*Contained by:* header: calendarDesc
*May contain*
*core:* p
*linking:* ab
*Example*

```
<calendarDesc>
  <calendar xml:id="julianEngland">
    <p>Julian Calendar (including proleptic)</p>
  </calendar>
</calendarDesc>
```

*Example*

```
<calendarDesc>
  <calendar xml:id="egyptian"
    target="http://en.wikipedia.org/wiki/Egyptian_calendar">
    <p>Egyptian calendar (as defined by Wikipedia)</p>
  </calendar>
</calendarDesc>
```

*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element calendar
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.pointing.attributes,
   att.typed.attributes,
   model.pLike+
}
```

---

## <calendarDesc>

<calendarDesc> (calendar description) contains a description of the calendar system used in any dating expression found in the text. [*2.4. The Profile Description 2.4.5. Calendar Description*]

*Module* header — The TEI Header
*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.profileDescPart

*Contained by*

*header:* profileDesc

*May contain*

*header:* calendar

*Note* In the first example above, calendars and short codes for *xml:id*s are from W3 guidelines at ??ttp://www.w3.org/TR/xpath-functions-11/#lang-cal-country

*Example*

```
<calendarDesc>
  <calendar  xml:id="cal_AD">
    <p>Anno Domini (Christian Era)</p>
  </calendar>
  <calendar  xml:id="cal_AH">
    <p>Anno Hegirae (Muhammedan Era)</p>
  </calendar>
  <calendar  xml:id="cal_AME">
    <p>Mauludi Era (solar years since Mohammed's birth)</p>
  </calendar>
```

```
<calendar xml:id="cal_AM">
  <p>Anno Mundi (Jewish Calendar)</p>
</calendar>
<calendar xml:id="cal_AP">
  <p>Anno Persici</p>
</calendar>
<calendar xml:id="cal_AS">
  <p>Aji Saka Era (Java)</p>
</calendar>
<calendar xml:id="cal_BE">
  <p>Buddhist Era</p>
</calendar>
<calendar xml:id="cal_CB">
  <p>Cooch Behar Era</p>
</calendar>
<calendar xml:id="cal_CE">
  <p>Common Era</p>
</calendar>
<calendar xml:id="cal_CL">
  <p>Chinese Lunar Era</p>
</calendar>
<calendar xml:id="cal_CS">
  <p>Chula Sakarat Era</p>
</calendar>
<calendar xml:id="cal_EE">
  <p>Ethiopian Era</p>
</calendar>
<calendar xml:id="cal_FE">
  <p>Fasli Era</p>
</calendar>
<calendar xml:id="cal_ISO">
  <p>ISO 8601 calendar</p>
</calendar>
<calendar xml:id="cal_JE">
  <p>Japanese Calendar</p>
</calendar>
<calendar xml:id="cal_KE">
  <p>Khalsa Era (Sikh calendar)</p>
</calendar>
<calendar xml:id="cal_KY">
  <p>Kali Yuga</p>
</calendar>
<calendar xml:id="cal_ME">
  <p>Malabar Era</p>
</calendar>
<calendar xml:id="cal_MS">
  <p>Monarchic Solar Era</p>
</calendar>
<calendar xml:id="cal_NS">
  <p>Nepal Samwat Era</p>
</calendar>
<calendar xml:id="cal_OS">
  <p>Old Style (Julian Calendar)</p>
</calendar>
<calendar xml:id="cal_RS">
  <p>Rattanakosin (Bangkok) Era</p>
</calendar>
```

```
  <calendar  xml:id="cal_SE">
    <p>Saka Era</p>
  </calendar>
  <calendar  xml:id="cal_SH">
    <p>Mohammedan Solar Era (Iran)</p>
  </calendar>
  <calendar  xml:id="cal_SS">
    <p>Saka Samvat</p>
  </calendar>
  <calendar  xml:id="cal_TE">
    <p>Tripurabda Era</p>
  </calendar>
  <calendar  xml:id="cal_VE">
    <p>Vikrama Era</p>
  </calendar>
  <calendar  xml:id="cal_VS">
    <p>Vikrama Samvat Era</p>
  </calendar>
</calendarDesc>
```

*Example*

```
<calendarDesc>
  <calendar  xml:id="cal_Gregorian">
    <p>Gregorian calendar</p>
  </calendar>
  <calendar  xml:id="cal_Julian">
    <p>Julian calendar</p>
  </calendar>
  <calendar  xml:id="cal_Islamic">
    <p>Islamic or Muslim (hijri) lunar calendar</p>
  </calendar>
  <calendar  xml:id="cal_Hebrew">
    <p>Hebrew or Jewish lunisolar calendar</p>
  </calendar>
  <calendar  xml:id="cal_Revolutionary">
    <p>French Revolutionary calendar</p>
  </calendar>
  <calendar  xml:id="cal_Iranian">
    <p>Iranian or Persian (Jalaali) solar calendar</p>
  </calendar>
  <calendar  xml:id="cal_Coptic">
    <p>Coptic or Alexandrian calendar</p>
  </calendar>
  <calendar  xml:id="cal_Chinese">
    <p>Chinese lunisolar calendar</p>
  </calendar>
</calendarDesc>
```

*Example*

```
<calendarDesc>
  <calendar  xml:id="cal_Egyptian"
    target="http://en.wikipedia.org/wiki/Egyptian_calendar">
    <p>Egyptian calendar (as defined by Wikipedia)</p>
  </calendar>
</calendarDesc>
```

*Content model*

```
<content>
 <elementRef key="calendar" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element calendarDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   calendar+
}
```

---

**<camera>** (camera) describes a particular camera angle or viewpoint in a screen play. [*7.3.1. Technical Information 7.3. Other Types of Performance Text*]

*Module* drama — Performance Texts

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.stageLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l lg meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound spGrp tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<view>George glances at the window—and freezes.
<camera  type="cut">New angle—shock cut</camera>
Out the window the body of a dead man suddenly slams into frame
</view>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element camera
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   macro.paraContent}
```

---

## <caption>

<caption> (caption) contains the text of a caption or other text displayed as part of a film script or screenplay. [*7.3.1. Technical Information 7.3. Other Types of Performance Text*]

*Module* drama — Performance Texts

*Attributes*

- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.stageLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l lg meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound spGrp tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* A specialized form of stage direction.

*Example*

```
<camera>Zoom in to overlay showing some stock film of hansom cabs
galloping past</camera>
<caption>London, 1895.</caption>
<caption>The residence of Mr Oscar Wilde.</caption>
<sound>Suitably classy music starts.</sound>
<view>Mix through to Wilde's drawing room. A crowd of suitably
dressed folk are engaged in typically brilliant conversation,
laughing affectedly and drinking champagne.</view>
<sp>
  <speaker>Prince of Wales</speaker>
  <p>My congratulations, Wilde. Your latest play is a great success.
  </p>
</sp>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element caption
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.paraContent}
```

---

**<case>** (case) contains grammatical case information given by a dictionary for a given form. [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*
  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.lexicographic.normalized
    * *@norm*
    * *@orig*

*Member of* model.morphLike

*Contained by*

*dictionaries:* dictScrap entryFree form gramGrp

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May contain character data and phrase-level elements. Typical values will be of the form *nominative*, *accusative*, *dative*, *genitive*, etc.

This element is synonymous with <gram type="case">.

*Example* Taken from *Wörterbuch der Deutschen Sprache. Veranstaltet und herausgegeben von Joachim Heinrich Campe. Erster Theil. A - bis - E. (Braunschweig 1807. In der Schulbuchhandlung)*:

Das Evangelium, des Evangelii, …

```
<entry>
  <form type="lemma">
    <gramGrp>
      <pos value="noun"/>
      <gen value="n"/>
    </gramGrp>
    <form type="determiner">
      <orth>Das</orth>
    </form>
    <form type="headword">
      <orth>Evangelium</orth>,</form>
  </form>
  <form type="inflected">
    <gramGrp>
      <case value="genitive"/>
      <number value="singular"/>
    </gramGrp>
    <form type="determiner">
      <orth>des</orth>
    </form>
    <form type="headword">
      <orth>
        <oRef>Evangelii</oRef>,</orth>
    </form>
  </form>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element case
{
```

```
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   macro.paraContent}
```

---

**<castGroup>** (cast list grouping) groups one or more individual <castItem> elements within a cast list. [*7.1.4. Cast Lists*]

*Module* drama — Performance Texts

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*

– att.global.source
* *@source*

*Contained by:* castGroup castList
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* cb ellipsis gap gb head index lb milestone note noteGrp pb
*drama:* castGroup castItem roleDesc
*figures:* figure notatedMusic
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*spoken:* incident kinesic pause shift vocal writing
*textcrit:* app witDetail
*textstructure:* trailer
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* The *rend* attribute may be used, as here, to indicate whether the grouping is indicated by a brace, whitespace, font change, etc.

Note that in this example the role description 'friends of Mathias' is understood to apply to both roles equally.

*Example*

```
<castGroup  rend="braced">
  <castItem>
    <role>Walter</role>
    <actor>Mr Frank Hall</actor>
  </castItem>
  <castItem>
    <role>Hans</role>
    <actor>Mr F.W. Irish</actor>
  </castItem>
  <roleDesc>friends of Mathias</roleDesc>
</castGroup>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.global"/>
   <classRef key="model.headLike"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <alternate>
    <elementRef key="castItem"/>
    <elementRef key="castGroup"/>
    <elementRef key="roleDesc"/>
   </alternate>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <sequence minOccurs="0">
```

```
   <elementRef key="trailer"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element castGroup
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      ( model.global | model.headLike )*,
      ( ( ( castItem | castGroup | roleDesc ), model.global* )+ ),
      ( ( trailer, model.global* )? )
   )
}
```

---

**<castItem>** (cast list item) contains a single entry within a cast list, describing either a single role or a list of non-speaking roles. [*7.1.4. Cast Lists*]

*Module* drama — Performance Texts

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

• att.typed
- ~~type~~
- @subtype

`@type` characterizes the cast item.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **role** the item describes a single role.[Default]
**list** the item describes a list of non-speaking roles.

*Contained by:* castGroup castList
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address binaryObject cb choice corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q ref reg rs ruby sic soCalled term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* actor role roleDesc
*figures:* figure formula notatedMusic
*gaiji:* g
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef specDesc specList tag val
*textcrit:* app witDetail
*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme
character data

*Example*

```
<castItem>
  <role>Player</role>
  <actor>Mr Milward</actor>
</castItem>
```

*Example*

```
<castItem  type="list">Constables, Drawer, Turnkey, etc.</castItem>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.castItemPart"/>
  <classRef key="model.phrase"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element castItem
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { "role" | "list" }?,
   ( text | model.gLike | model.castItemPart | model.phrase | model.global )*
}
```

---

## <castList>

(cast list) contains a single cast list or dramatis personae. [*7.1.4. Cast Lists 7.1. Front and Back Matter* ]

*Module* drama — Performance Texts

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.frontPart.drama model.inter model.standOffPart

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph front imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct cb cit desc ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castGroup castItem castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline dateline docAuthor docDate epigraph floatingText opener salute signed

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Example*

```
<castList>
  <castGroup>
    <head  rend="braced">Mendicants</head>
    <castItem>
      <role>Aafaa</role>
      <actor>Femi Johnson</actor>
    </castItem>
    <castItem>
      <role>Blindman</role>
      <actor>Femi Osofisan</actor>
    </castItem>
    <castItem>
      <role>Goyi</role>
      <actor>Wale Ogunyemi</actor>
    </castItem>
    <castItem>
      <role>Cripple</role>
      <actor>Tunji Oyelana</actor>
    </castItem>
  </castGroup>
  <castItem>
    <role>Si Bero</role>
    <roleDesc>Sister to Dr Bero</roleDesc>
    <actor>Deolo Adedoyin</actor>
  </castItem>
  <castGroup>
    <head  rend="braced">Two old women</head>
    <castItem>
      <role>Iya Agba</role>
      <actor>Nguba Agolia</actor>
```

```
      </castItem>
      <castItem>
        <role>Iya Mate</role>
        <actor>Bopo George</actor>
      </castItem>
    </castGroup>
    <castItem>
      <role>Dr Bero</role>
      <roleDesc>Specialist</roleDesc>
      <actor>Nat Okoro</actor>
    </castItem>
    <castItem>
      <role>Priest</role>
      <actor>Gbenga Sonuga</actor>
    </castItem>
    <castItem>
      <role>The old man</role>
      <roleDesc>Bero's father</roleDesc>
      <actor>Dapo Adelugba</actor>
    </castItem>
  </castList>
  <stage  type="mix">The action takes place in and around the home surgery of
  Dr Bero, lately returned from the wars.</stage>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.common"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <alternate>
    <elementRef key="castItem"/>
    <elementRef key="castGroup"/>
   </alternate>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.common"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element castList
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      ( model.divTop | model.global )*,
      ( ( model.common, model.global* )* ),
      ( ( ( castItem | castGroup ), model.global* )+ ),
      ( ( model.common, model.global* )* )
   )
}
```

---

**<catchwords>** (catchwords) describes the system used to ensure correct ordering of the quires or similar making up a codex, incunable, or other object typically by means of annotations at the foot of the page. [*11.3.7. Catchwords, Signatures, Secundo Folio*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<catchwords>Vertical catchwords in the hand of the scribe placed along
the inner bounding line, reading from top to bottom.</catchwords>
```

*Schematron* <sch:rule context="tei:catchwords"> <sch:assert test="ancestor::tei:msDesc or ancestor::tei:egXML">The <sch:name/> element should not be used outside of msDesc.</sch:assert> </sch:rule>

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element catchwords
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

---

## <catDesc>

<catDesc> (category description) describes some category within a taxonomy or text typology, either in the form of a brief prose description or in terms of the situational parameters used by the TEI formal <textDesc>. [*2.3.7. The Classification Declaration*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.canonical
  - *@key*
  - *@ref*

*Contained by:* category

*May contain*

*core:* abbr address choice date distinct email emph expan foreign gloss hi measure measureGrp mentioned name num ptr q ref rs soCalled term time title unit

*corpus:* textDesc

*dictionaries:* lang

*header:* idno

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*tagdocs:* att code gi ident listRef tag val

*transcr:* am ex subst

character data

*Example*

```
<catDesc>Prose reportage</catDesc>
```

*Example*

```
<catDesc>
  <textDesc n="novel">
    <channel mode="w">print; part issues</channel>
    <constitution type="single"/>
    <derivation type="original"/>
    <domain type="art"/>
    <factuality type="fiction"/>
    <interaction type="none"/>
    <preparedness type="prepared"/>
    <purpose type="entertain" degree="high"/>
    <purpose type="inform" degree="medium"/>
  </textDesc>
</catDesc>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.limitedPhrase"/>
  <classRef key="model.catDescPart"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element catDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   ( text | model.limitedPhrase | model.catDescPart )*
}
```

**<category>** (category) contains an individual descriptive category, possibly nested within a superordinate category, within a user-defined taxonomy. [*2.3.7. The Classification Declaration*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*

*Contained by* *header:* category taxonomy

*May contain*

*core:* desc gloss

*header:* catDesc category

*tagdocs:* equiv

*Example*

```
<category xml:id="b1">
  <catDesc>Prose reportage</catDesc>
</category>
```

*Example*

```
<category xml:id="b2">
  <catDesc>Prose </catDesc>
  <category xml:id="b11">
    <catDesc>journalism</catDesc>
  </category>
  <category xml:id="b12">
    <catDesc>fiction</catDesc>
  </category>
</category>
```

*Example*

```
<category xml:id="LIT">
  <catDesc xml:lang="pl">literatura piękna</catDesc>
  <catDesc xml:lang="en">fiction</catDesc>
  <category xml:id="LPROSE">
    <catDesc xml:lang="pl">proza</catDesc>
    <catDesc xml:lang="en">prose</catDesc>
  </category>
  <category xml:id="LPOETRY">
    <catDesc xml:lang="pl">poezja</catDesc>
    <catDesc xml:lang="en">poetry</catDesc>
  </category>
  <category xml:id="LDRAMA">
    <catDesc xml:lang="pl">dramat</catDesc>
    <catDesc xml:lang="en">drama</catDesc>
  </category>
</category>
```

*Content model*

```
<content>
 <sequence>
  <alternate>
   <elementRef key="catDesc" minOccurs="1"
    maxOccurs="unbounded"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.descLike"/>
    <elementRef key="equiv"/>
    <elementRef key="gloss"/>
   </alternate>
  </alternate>
  <elementRef key="category" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element category
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datcat.attributes,
   ( ( catDesc+ | ( model.descLike | equiv | gloss )* ), category* )
}
```

---

**<catRef>** (category reference) specifies one or more defined categories within some taxonomy or text typology. [*2.4.3. The Text Classification*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

  * *@cert*
  * *@resp*
- att.global.source
  * *@source*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*

`@scheme` identifies the classification scheme within which the set of categories concerned is defined, for example by a <taxonomy> element, or by some other resource.
*Status* Optional
*Datatype* teidata.pointer

*Contained by* imprint
*header:* textClass
*May contain* Empty element
*Note* The *scheme* attribute needs to be supplied only if more than one taxonomy has been declared.
*Example*

```
<catRef  scheme="#myTopics"
  target="#news #prov #sales2"/>
<!-- elsewhere -->
<taxonomy  xml:id="myTopics">
  <category  xml:id="news">
    <catDesc>Newspapers</catDesc>
  </category>
  <category  xml:id="prov">
    <catDesc>Provincial</catDesc>
  </category>
  <category  xml:id="sales2">
    <catDesc>Low to average annual sales</catDesc>
  </category>
</taxonomy>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element catRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.pointing.attributes,
   attribute scheme { text }?,
   empty
}
```

**<cb>** (column beginning) marks the beginning of a new column of a text on a multi-column page. [*3.11.3. Milestone Elements*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.breaking
  - *@break*
- att.cmc
  - *@generatedBy*
- att.edition
  - *@ed*
  - *@edRef*
- att.spanning

- *@spanTo*

- att.typed
  - *@type*
  - *@subtype*

*Member of* model.milestoneLike

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list listBibl measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc subst supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Note* On this element, the global *n* attribute indicates the number or other value associated with the column which follows the point of insertion of this <cb> element. Encoders should adopt a clear and consistent policy as to whether the numbers associated with column beginnings relate to the physical sequence number of the column in the whole text, or whether columns are numbered within the page. The <cb> element is placed at the head of the column to which it refers.

*Example* Markup of an early English dictionary printed in two columns:

```
<pb/>
<cb  n="1"/>
```

```
<entryFree>
  <form>Well</form>, <sense>a Pit to hold Spring-Water</sense>:
<sense>In the Art of <hi  rend="italic">War</hi>, a Depth the Miner
    sinks into the Ground, to find out and disappoint the Enemies Mines,
    or to prepare one</sense>.
</entryFree>
<entryFree>To <form>Welter</form>, <sense>to wallow</sense>, or
<sense>lie groveling</sense>.</entryFree>
<!-- remainder of column -->
<cb  n="2"/>
<entryFree>
  <form>Wey</form>, <sense>the greatest Measure for dry Things,
    containing five Chaldron</sense>.
</entryFree>
<entryFree>
  <form>Whale</form>, <sense>the greatest of
    Sea-Fishes</sense>.
</entryFree>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element cb
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.breaking.attributes,
   att.cmc.attributes,
   att.edition.attributes,
   att.spanning.attributes,
   att.typed.attributes,
   empty
}
```

---

## <cell>

(cell) contains one cell of a table. [*15.1.1. TEI Tables*]

*Module* figures — Tables, Formulæ, Graphics, and Notated Music

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
- att.global.analytic
  - *@ana*

  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.tableDecoration
  - *@role*
  - *@rows*
  - *@cols*

*Contained by:* figure row
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound spGrp tech view
*figures:* figure formula notatedMusic table
*gaiji:* g
*header:* biblFull idno
*iso-fs:* fLib fs fvLib
*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat

geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<row>
  <cell  role="label">General conduct</cell>
  <cell  role="data">Not satisfactory, on account of his great unpunctuality
     and inattention to duties</cell>
</row>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element cell
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.tableDecoration.attributes,
   macro.specialPara}
```

---

**<certainty>** indicates the degree of certainty associated with some aspect of the text markup. [*22.1.2. Structured Indications of Uncertainty*]

*Module* certainty — Certainty, Precision, and Responsibility

*Attributes*

- att.scoping
  - *@target*
  - *@match*
- att.global
  - @xml:id

- @n
- @xml:lang
- @xml:base
- @xml:space

- att.global.analytic
  - @ana
- att.global.change
  - @change
- att.global.facs
  - @facs
- att.global.linking
  - @corresp
  - @synch
  - @sameAs
  - @copyOf
  - @next
  - @prev
  - @exclude
  - @select
- att.global.rendition
  - @rend
  - @style
  - @rendition
- att.global.responsibility
  - ~~cert~~
  - @resp
- att.global.source
  - @source
- att.typed
  - ~~type~~
  - @subtype

`@cert` (certainty) signifies the degree of certainty associated with the object pointed to by the <certainty> element.
*Derived from* att.global.responsibility
*Status* Optional
*Datatype* teidata.certainty

`@type` characterizes the element in some sense, using any convenient classification scheme or typology; sample categorization of annotations of uncertainty might use following values:
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **ignorance**
**incompleteness**
**credibility**

**imprecision**

`@locus` indicates more exactly the aspect concerning which certainty is being expressed: specifically, whether the markup is correctly located, whether the correct element or attribute name has been used, or whether the content of the element or attribute is correct, etc.

*Status* Required

*Datatype* teidata.enumerated

*Legal values are:* **name** uncertainty concerns whether the name of the element or attribute used is correctly applied.

**start** uncertainty concerns whether the start of the element is correctly identified.

**end** uncertainty concerns whether the end of the element is correctly identified.

**location** uncertainty concerns both the start and the end of the element.

**value** uncertainty concerns the content (for an element) or the value (for an attribute)

`@assertedValue` provides an alternative value for the aspect of the markup in question—an alternative generic identifier, transcription, or attribute value, or the identifier of an <anchor> element (to indicate an alternative starting or ending location). If an *assertedValue* is given, the confidence level specified by *degree* applies to the alternative markup specified by *assertedValue*; if none is given, it applies to the markup in the text.

*Status* Recommended

*Datatype* teidata.word

```
<certainty target="#ESX"
  locus="name"
  assertedValue="placeName"
  cert="low">
  <desc>It is unlikely, but possible, that this refers to the
place
    rather than the person.</desc>
</certainty>
```

*Note* This attribute makes it possible to indicate the degree of confidence in a specific alternative to some aspect of the markup. In the first example above the encoder is expressing the likelihood (.2) that the generic identifier should be <placeName> rather than <persName>, which is the coded element.

`@given` indicates one or more element(s) characterizing the conditions which are assumed in the assignment of a degree of confidence.

*Status* Recommended

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Note* The envisioned typical value of this attribute would be the identifier of another <certainty> element or a list of such identifiers. It may thus be possible to construct probability networks by chaining <certainty> elements together. Such networks would ultimately be grounded in unconditional <certainty> elements (with no value for *given*). The semantics of this chaining would be understood in this way: if a <certainty> element is specified, via a reference, as the assumption, then it is not the attribution of uncertainty that is the assumption, but rather the assertion itself. For instance, in the example above, the first <certainty> element indicates that the confidence in the identification of the new scribe as *msm*. The second indicates the degree of confidence that *Essex* is a personal name, given that the new scribe is *msm*. Note that the given in the second <certainty> element is not the assertion that the likelihood that msm is the new scribe is 0.6, but simply the assertion that msm is the new scribe; this is a recommended convention to facilitate building networks.

The ambitious encoder may wish to attempt complex networks or probability assertions, experimenting with references to other elements or prose assertions, and

deploying feature structure connectives such as <alt>, <join>, and <note>. However, we do not believe that the <certainty> element gives, at this time, a comprehensive ambiguity-free system for indicating certainty.

`@degree` indicates the degree of confidence assigned to the aspect of the markup named by the *locus* attribute.
*Status* Optional
*Datatype* teidata.probability

*Member of* model.certLike model.global.meta

*Contained by*

*analysis:* cl interp m phr s span w

*certainty:* certainty precision respons

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gap gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab join seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc space substJoin supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*certainty:* certainty precision respons

*core:* desc

*Example* (For discussion of this example, see section *22.1.2.1. Contingent Conditions*)

```
Ernest went to <anchor  xml:id="A1"/> old
<persName  xml:id="SYB">Saybrook</persName>.

<certainty  xml:id="c1"  target="#SYB"
  locus="name"  degree="0.6"/>
<certainty  target="#SYB"  locus="start"
  given="#c1"  degree="0.9"/>
<certainty  xml:id="C-c2"  target="#SYB"
  locus="name"  assertedValue="persName"  degree="0.4"/>
<certainty  target="#SYB"  locus="start"
  given="#C-c2"  degree="0.5"/>
<certainty  target="#SYB"  locus="start"
  assertedValue="#a1"  given="#c1"  degree="0.5"/>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.descLike"/>
  <classRef key="model.certLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element certainty
{
   att.scoping.attributes,
   att.global.attribute.xmlid,
   att.global.attribute.n,
   att.global.attribute.xmllang,
   att.global.attribute.xmlbase,
   att.global.attribute.xmlspace,
   att.global.analytic.attribute.ana,
   att.global.change.attribute.change,
   att.global.facs.attribute.facs,
   att.global.linking.attribute.corresp,
   att.global.linking.attribute.synch,
   att.global.linking.attribute.sameAs,
   att.global.linking.attribute.copyOf,
   att.global.linking.attribute.next,
   att.global.linking.attribute.prev,
   att.global.linking.attribute.exclude,
   att.global.linking.attribute.select,
   att.global.rendition.attribute.rend,
   att.global.rendition.attribute.style,
   att.global.rendition.attribute.rendition,
   att.global.responsibility.attribute.resp,
   att.global.source.attribute.source,
   att.typed.attribute.subtype,
   attribute cert { text }?,
   attribute type { text }?,
   attribute locus { "name" | "start" | "end" | "location" | "value" },
   attribute assertedValue { text }?,
   attribute given { list { + } }?,
   attribute degree { text }?,
```

```
   ( model.descLike | model.certLike )*
}
```

---

**<change>** (change) documents a change or set of changes made during the production of a source document, or during the revision of an electronic file. [*2.6. The Revision Description 2.4.1. Creation 12.7. Identifying Changes and Revisions*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.ascribed
  - *@who*
- att.datable
  - *@period*

- att.datable.custom
  - *@when-custom*
  - *@notBefore-custom*
  - *@notAfter-custom*
  - *@from-custom*
  - *@to-custom*
  - *@datingPoint*
  - *@datingMethod*
- att.datable.iso
  - *@when-iso*
  - *@notBefore-iso*
  - *@notAfter-iso*
  - *@from-iso*
  - *@to-iso*
- att.datable.w3c
  - *@when*
  - *@notBefore*
  - *@notAfter*
  - *@from*
  - *@to*
- att.docStatus
  - *@status*
- att.typed
  - *@type*
  - *@subtype*

`@target` (target) points to one or more elements that belong to this change.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Contained by:* *header:* listChange revisionDesc

*msdescription:* recordHist

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The *who* attribute may be used to point to any other element, but will typically specify a <respStmt> or <person> element elsewhere in the header, identifying the person responsible for the change and their role in making it.

It is recommended that changes be recorded with the most recent first. The *status* attribute may be used to indicate the status of a document following the change documented.

*Example*

```
<titleStmt>
  <title> ... </title>
  <editor  xml:id="LDB">Lou Burnard</editor>
  <respStmt  xml:id="BZ">
    <resp>copy editing</resp>
    <name>Brett Zamir</name>
  </respStmt>
</titleStmt>
<!-- ... -->
<revisionDesc  status="published">
  <change  who="#BZ"  when="2008-02-02"
    status="public">Finished chapter 23</change>
  <change  who="#BZ"  when="2008-01-02"
    status="draft">Finished chapter 2</change>
  <change  n="P2.2"  when="1991-12-21"
    who="#LDB">Added examples to section 3</change>
  <change  when="1991-11-11"  who="#MSM">Deleted chapter 10</change>
</revisionDesc>
```

*Example*

```
<profileDesc>
  <creation>
    <listChange>
      <change  xml:id="DRAFT1">First draft in pencil</change>
      <change  xml:id="DRAFT2"
        notBefore="1880-12-09">First revision, mostly
            using green ink</change>
```

```
          <change  xml:id="DRAFT3"
            notBefore="1881-02-13">Final corrections as
                supplied to printer.</change>
        </listChange>
      </creation>
    </profileDesc>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element change
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.docStatus.attributes,
   att.typed.attributes,
   attribute target { list { + } }?,
   macro.specialPara}
```

---

## <channel>

<channel> (primary channel) describes the medium or channel by which a text is delivered or experienced. For a written text, this might be print, manuscript, email, etc.; for a spoken one, radio, telephone, face-to-face, etc. [*16.2.1. The Text Description*]

*Module* corpus — Language Corpora

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@mode` specifies the mode of this channel with respect to speech and writing.
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **s** (spoken)
**w** (written)
**sw** (spoken to be written) e.g. dictation
**ws** (written to be spoken) e.g. a script
**m** (mixed)
**x** (unknown or inapplicable) [Default]

*Member of* model.textDescPart
*Contained by*
*corpus:* textDesc
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit
*dictionaries:* lang
*figures:* figure notatedMusic
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data

*Example*

```
<channel mode="s">face-to-face conversation</channel>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element channel
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute mode { "s" | "w" | "sw" | "ws" | "m" | "x" }?,
   macro.phraseSeq.limited}
```

---

## <char>

(character) provides descriptive information about a character. [*5.2. Markup Constructs for Representation of Characters and Glyphs*]

*Module* gaiji — Characters, Glyphs, and Writing Modes

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*

- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Contained by* *gaiji:* charDecl
*May contain*
*core:* binaryObject desc graphic media note noteGrp
*figures:* figure formula
*gaiji:* localProp mapping unicodeProp unihanProp
*Example*

```
<char xml:id="circledU4EBA">
  <localProp name="Name"
    value="CIRCLED IDEOGRAPH 4EBA"/>
  <localProp name="daikanwa" value="36"/>
  <unicodeProp name="Decomposition_Mapping"
    value="circle"/>
  <mapping type="standard">人</mapping>
</char>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <elementRef key="unicodeProp"/>
  <elementRef key="unihanProp"/>
  <elementRef key="localProp"/>
  <elementRef key="mapping"/>
  <elementRef key="figure"/>
  <classRef key="model.graphicLike"/>
  <classRef key="model.noteLike"/>
  <classRef key="model.descLike"/>
 </alternate>
```

```
</content>
```

*Schema Declaration*

```
element char
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      unicodeProp    | unihanProp    | localProp    | mapping    | figure    | model.graph
}
```

---

## <charDecl>

(character declarations) provides information about nonstandard characters and glyphs. [*5.2. Markup Constructs for Representation of Characters and Glyphs*]

*Module* gaiji — Characters, Glyphs, and Writing Modes

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.encodingDescPart
*Contained by*
*header:* encodingDesc
*May contain*
*core:* desc
*gaiji:* char glyph
*Example*

```
<charDecl>
  <char xml:id="aENL">
    <unicodeProp name="Name"
      value="LATIN LETTER ENLARGED SMALL A"/>
    <mapping type="standard">a</mapping>
  </char>
</charDecl>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="desc" minOccurs="0"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="char"/>
   <elementRef key="glyph"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element charDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( desc?, ( char | glyph )+ )
}
```

**<choice>** (choice) groups a number of alternative encodings for the same point in a text. [*3.5. Simple Editorial Changes*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*

*Member of* model.linePart model.pPart.editorial

*Contained by*

*analysis:* cl pc phr s span w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* metSym rhyme

*May contain*

*core:* abbr choice corr expan orig reg sic unclear

*linking:* seg

*transcr:* am ex supplied

*Note* Because the children of a <choice> element all represent alternative ways of encoding the same sequence, it is natural to think of them as mutually exclusive. However, there may be cases where a full representation of a text requires the alternative encodings to be considered as parallel.

Note also that <choice> elements may self-nest.

Where the purpose of an encoding is to record multiple witnesses of a single work, rather than to identify multiple possible encoding decisions at a given point, the <app> element and associated elements discussed in section *13.1. The Apparatus Entry, Readings, and Witnesses* should be preferred.

*Example* An American encoding of *Gulliver's Travels* which retains the British spelling but also provides a version regularized to American spelling might be encoded as follows.

```
<p>Lastly, That, upon his solemn oath to observe all the above
articles, the said man-mountain shall have a daily allowance of
meat and drink sufficient for the support of <choice>
   <sic>1724</sic>
   <corr>1728</corr>
  </choice> of our subjects,
with free access to our royal person, and other marks of our
<choice>
   <orig>favour</orig>
   <reg>favor</reg>
  </choice>.</p>
```

*Content model*

```
<content>
```

```
 <alternate minOccurs="2"
  maxOccurs="unbounded">
  <classRef key="model.choicePart"/>
  <elementRef key="choice"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element choice
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   ( model.choicePart | choice ),
   ( model.choicePart | choice ),
   ( model.choicePart | choice )*
}
```

---

**<cit>** (cited quotation) contains a quotation from some other document, together with a bibliographic reference to its source. In a dictionary it may contain an example text with at least one occurrence of the word form, used in the sense being described, or a translation of the headword, or an example. [*3.3.3. Quotation 4.3.1. Grouped Texts 10.3.5.1. Examples*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*

    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.entryPart.top model.quoteLike

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author biblScope cit citedRange corr del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled sp speaker stage street term textLang title unclear unit

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castList epilogue performance prologue role roleDesc set sound tech view

*figures:* cell figDesc figure

*header:* change distributor edition extent geoDecl handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material msItem musicNotation objectType origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* interp interpGrp pc span spanGrp

*certainty:* certainty precision respons

*core:* bibl biblStruct binaryObject cb cit ellipsis gap gb graphic index lb listBibl media milestone note noteGrp pb ptr q quote ref said

*dictionaries:* colloc def etym form gramGrp hom hyph lbl orth pos pron re sense subc superEntry syll usg xr

*figures:* figure formula notatedMusic

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* eg egXML listRef

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Example*

```
<cit>
  <quote>and the breath of the whale is frequently attended with such an
insupportable smell,
    as to bring on disorder of the brain.</quote>
  <bibl>Ulloa's South America</bibl>
</cit>
```

*Example*

```
<entry>
  <form>
    <orth>horrifier</orth>
  </form>
  <cit  type="translation"  xml:lang="en">
    <quote>to horrify</quote>
  </cit>
  <cit  type="example">
    <quote>elle était horrifiée par la dépense</quote>
    <cit  type="translation"  xml:lang="en">
      <quote>she was horrified at the expense.</quote>
    </cit>
  </cit>
</entry>
```

*Example*

```
<cit  type="example">
  <quote  xml:lang="mix">Ka'an yu tsa'a Pedro.</quote>
  <media  url="soundfiles-gen:S_speak_1s_on_behalf_of_Pedro_01_02_03_TS.wav"
    mimeType="audio/wav"/>
  <cit  type="translation">
    <quote  xml:lang="en">I'm speaking on behalf of Pedro.</quote>
  </cit>
  <cit  type="translation">
```

```
      <quote  xml:lang="es">Estoy hablando de parte de Pedro.</quote>
    </cit>
  </cit>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.biblLike"/>
  <classRef key="model.egLike"/>
  <classRef key="model.entryPart"/>
  <classRef key="model.global"/>
  <classRef key="model.graphicLike"/>
  <classRef key="model.ptrLike"/>
  <classRef key="model.attributable"/>
  <elementRef key="pc"/>
  <elementRef key="q"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element cit
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.typed.attributes,
   (
      model.biblLike    | model.egLike    | model.entryPart    | model.global    | model.g
}
```

---

## <citeData>

(citation data) specifies how information may be extracted from citation structures. [*3.11.4. Declaring Reference Systems 17.2.5.4. Citation Structures*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.citeStructurePart
  - *@use*

`@property` (property) A URI indicating a property definition.
*Status* Required
*Datatype* teidata.pointer

*Contained by:* citeStructure
*May contain* Empty element
*Example*

```
<citeStructure unit="book"
  match="//body/div" use="@n">
  <citeData property="http://purl.org/dc/terms/title"
    use="head"/>
</citeStructure>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element citeData
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
```

```
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.citeStructurePart.attributes,
   attribute property { text },
   empty
}
```

---

**<citedRange>** (cited range) defines the range of cited content, often represented by pages or other units.
[*3.12.2.5. Scopes and Ranges in Bibliographic Citations*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.citing

  - *@unit*
  - *@from*
  - *@to*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*

*Member of* model.biblPart

*Contained by*

*core:* bibl biblStruct

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* When a single page is being cited, use the *from* and *to* attributes with an identical value. When no clear endpoint is provided, the *from* attribute may be used without *to*; for example a citation such as 'p. 3ff' might be encoded `<citedRange from="3">p. 3ff</citedRange>`.

*Example*

```
<citedRange>pp 12–13</citedRange>
<citedRange unit="page" from="12" to="13"/>
<citedRange unit="volume">II</citedRange>
<citedRange unit="page">12</citedRange>
```

*Example*

```
<bibl>
  <ptr  target="#mueller01"/>,
<citedRange  target="http://example.com/mueller3.xml#page4">vol. 3, pp.
    4-5</citedRange>
</bibl>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element citedRange
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.citing.attributes,
   att.pointing.attributes,
   macro.phraseSeq}
```

---

**<citeStructure>** (citation structure) declares a structure and method for citing the current document. [*3.11.4. Declaring Reference Systems 17.2.5.4. Citation Structures*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

- att.citeStructurePart
  - *@use*

`@delim` (delimiter) supplies a delimiting string preceding the structural component.

*Status* Optional

*Datatype* string

*Note* *delim* must contain at least one character.

`@match` (match) supplies an XPath selection pattern using the syntax defined in **?** which identifies a set of nodes which are citable structural components. The expression may be absolute (beginning with `/`) or relative. *match* on a <citeStructure> without a <citeStructure> parent must be an absolute XPath. If it is relative, its context is set by the *match* of the parent <citeStructure>.

*Status* Required

*Datatype* teidata.xpath

*Schematron* <sch:rule context="tei:citeStructure[not(parent::tei:citeStructure)]"> <sch:assert test="starts-with(@match,'/')">An XPath in @match on the outer <sch:name/> must start with '/'.</sch:assert> </sch:rule>

*Schematron* <sch:rule context="tei:citeStructure[parent::tei:citeStructure]"> <sch:assert test="not(starts-with(@match,'/'))">An XPath in @match must not start with '/' except on the outer <sch:name/>.</sch:assert> </sch:rule>

`@unit` (unit) describes the structural unit indicated by the <citeStructure>.

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **book**

**chapter**

**entry**

**poem**

**letter**

**line**

**section**

**verse**

**volume**

*Contained by* *header:* citeStructure refsDecl
*May contain*
*core:* desc
*header:* citeData citeStructure
*Example*

```
<citeStructure unit="book"
  match="//body/div" use="@n">
  <citeStructure unit="chapter" match="div"
    use="position()" delim=" ">
    <citeStructure unit="verse" match="div"
      use="position()" delim=":"/>
  </citeStructure>
</citeStructure>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="citeData" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="citeStructure"
   minOccurs="0" maxOccurs="unbounded"/>
  <classRef key="model.descLike"
   minOccurs="0" maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element citeStructure
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.citeStructurePart.attributes,
   attribute delim { text }?,
   attribute match { text },
   attribute unit { text }?,
   ( citeData*, citeStructure*, model.descLike* )
}
```

---

**<cl>** (clause) represents a grammatical clause. [*18.1. Linguistic Segment Categories*]
*Module* analysis — Simple Analytic Mechanisms
*Attributes* • att.global
- *@xml:id*
- *@n*

  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.notated
  - *@notation*
- att.segLike
  - *@function*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.fragmentable
    * *@part*
  - att.metrical
    * *@met*

    - *@real*
    - *@rhyme*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.segLike

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The *type* attribute may be used to indicate the type of clause, taking values such as finite, nonfinite, declarative, interrogative, relative etc. as appropriate.

*Example*

```
<l>
  <cl  type="relative"
    function="clause_modifier">Which frightened
     both the heroes so,</cl>
</l>
<l>
  <cl>They quite forgot their quarrel.</cl>
</l>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element cl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.notated.attributes,
   att.segLike.attributes,
   att.datcat.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

**<classCode>** (classification code) contains the classification code used for this text in some standard classification system. [*2.4.3. The Text Classification*]

*Module* header — The TEI Header

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@scheme` identifies the classification system in use, as defined by, e.g. a <taxonomy> element, or some other resource.
*Status* Required
*Datatype* teidata.pointer

*Contained by:* imprint
*header:* textClass
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin

character data

*Example*

```
<classCode scheme="http://www.udc.org">410</classCode>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element classCode
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute scheme { text },
   macro.phraseSeq.limited}
```

---

## <classDecl>

(classification declarations) contains one or more taxonomies defining any classificatory codes used elsewhere in the text. [*2.3.7. The Classification Declaration 2.3. The Encoding Description*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.encodingDescPart

*Contained by*

*header:* encodingDesc

*May contain*

*header:* taxonomy

*Example*

```
<classDecl>
  <taxonomy xml:id="LCSH">
    <bibl>Library of Congress Subject Headings</bibl>
  </taxonomy>
</classDecl>
<!-- ... -->
<textClass>
  <keywords scheme="#LCSH">
    <term>Political science</term>
    <term>United States -- Politics and government --
        Revolution, 1775-1783</term>
  </keywords>
</textClass>
```

*Content model*

```
<content>
 <elementRef key="taxonomy" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element classDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   taxonomy+
}
```

---

## <classes>

(classes) specifies all the classes of which the documented element or class is a member or subclass.
[*23.5. Element Specifications 23.6. Class Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*
– att.global.responsibility
* *@cert*
* *@resp*
– att.global.source
* *@source*

`@mode` specifies the effect of this declaration on its parent module.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **change** this declaration changes the declaration of the same name in the current definition

**replace** this declaration replaces the declaration of the same name in the current definition[Default]

*Contained by tagdocs:* classSpec elementSpec

*May contain*

*tagdocs:* memberOf

*Note* An empty <classes> element indicates that the element documented is not a member of any class. This should not generally happen.

*Example*

```
<classes>
  <memberOf  key="model.attributable"/>
  <memberOf  key="att.declarable"/>
</classes>
```

This <classes> element indicates that the element documented (which may be an element or a class) is a member of two distinct classes: `model.attributable` and `att.declarable`.

*Content model*

```
<content>
 <elementRef key="memberOf" minOccurs="0"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element classes
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute mode { "change" | "replace" }?,
```

```
    memberOf*
}
```

---

**<classRef>** points to the specification for an attribute or model class which is to be included in a schema. [*23.6. Class Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.repeatable
  - *@minOccurs*
  - *@maxOccurs*

`@key` the identifier used for the required class within the source indicated.

*Status* Required

*Datatype* teidata.xmlName

`@expand` indicates how references to this class within a content model should be interpreted.

*Status* Optional

*Legal values are:* **alternation** any one member of the class may appear

**sequence** a single occurrence of all members of the class may appear in sequence

**sequenceOptional** a single occurrence of one or more members of the class may appear in sequence

**sequenceOptionalRepeatable** one or more occurrences of one or more members of the class may appear in sequence.

**sequenceRepeatable** one or more occurrences of all members of the class may appear in sequence

*Note* If the members of the class are `a`, `b` and `c`, then a reference to the class within a content model is understood as being a reference to `a|b|c` when *expand* is omitted or has the value alternation; to `a,b,c` when it has the value sequence; to `(a?,b?,c?)` when it has the value sequenceOptional; to `(a*,b*,c*)` when it has the value sequenceOptionalRepeatable; or to `(a+,b+,c+)` when it has the value sequenceRepeatable.

`@include` supplies a list of class members which are to be included in the schema being defined.

*Status* Optional

*Datatype* 0–∞ occurrences of teidata.xmlName separated by whitespace

`@except` supplies a list of class members which are to be excluded from the schema being defined.

*Status* Optional

*Datatype* 0–∞ occurrences of teidata.xmlName separated by whitespace

*Member of* model.contentPart model.oddRef

*Contained by*

*tagdocs:* alternate content schemaSpec sequence specGrp

*May contain* Empty element

*Note* Attribute and model classes are identified by the name supplied as value for the *ident* attribute on the <classSpec> element in which they are declared. All TEI class names are unique; attribute class names conventionally begin with the letters `att.`.

*Example*

```
<schemaSpec ident="myTEIe" start="TEI">
  <moduleRef key="tei"/>
  <moduleRef key="core"
    include="abbr add core addrLine address author bibl biblScope choice cit
corr date del desc divGen editor emph expan foreign gap gloss graphic head hi
index item l label lb lg list listBibl mentioned milestone name note num orig p
pb ptr pubPlace publisher q ref reg relatedItem resp respStmt rs sic soCalled
sp speaker stage teiCorpus term time title unclear"/>
  <moduleRef key="header"
    include="authority availability catRef category change classCode classDecl
creation distributor edition editionStmt editorialDecl encodingDesc extent
fileDesc funder idno keywords langUsage language licence notesStmt principal
profileDesc projectDesc publicationStmt refsDecl revisionDesc samplingDecl
seriesStmt sourceDesc sponsor taxonomy teiHeader textClass titleStmt"/>
  <moduleRef key="textstructure"
    include="TEI argument back body byline closer dateline div docAuthor
docDate docEdition docImprint docTitle epigraph front group opener postscript
salute signed text titlePage titlePart trailer"/>
  <classRef key="att.global.facs"/>
</schemaSpec>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element classRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.repeatable.attributes,
   attribute key { text },
   attribute expand
   {
      "alternation"
    | "sequence"
    | "sequenceOptional"
    | "sequenceOptionalRepeatable"
    | "sequenceRepeatable"
   }?,
   ( attribute include { list { * } }? | attribute except { list { * } }? ),
   empty
}
```

---

## <classSpec>

**<classSpec>** (class specification) contains reference information for a TEI element class; that is a group of elements which appear together in content models, or which share some common attribute, or both. [*23.3. Specification Elements 23.6. Class Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*

  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

- att.identified
  - *@ident*
  - *@predeclare*
  - *@module*

  - att.combinable
    * *@mode*

    * att.deprecated
    · *@validUntil*

- att.typed
  - ~~type~~
  - @subtype

`@type` indicates whether this is a model class or an attribute class
*Derived from* att.typed
*Status* Required
*Datatype* teidata.enumerated
*Legal values are:* **model** (content model) members of this class appear in the same content models
**atts** (attributes) members of this class share common attributes

`@generate` indicates which alternation and sequence instantiations of a model class may be referenced. By default, all variations are permitted.
*Status* Optional
*Datatype* 1–5 occurrences of teidata.enumerated separated by whitespace
*Legal values are:* **alternation** members of the class are alternatives
**sequence** members of the class are to be provided in sequence
**sequenceOptional** members of the class may be provided, in sequence, but are optional
**sequenceOptionalRepeatable** members of the class may be provided one or more times, in sequence, but are optional.
**sequenceRepeatable** members of the class may be provided one or more times, in sequence

*Member of* model.oddDecl

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*tagdocs:* schemaSpec specGrp

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc gloss

*tagdocs:* altIdent attList classes constraintSpec equiv exemplum listRef remarks

*Example*

```
<classSpec  module="tei"  type="model"
  ident="model.segLike">
  <desc>groups elements used for arbitrary segmentation.</desc>
  <classes>
    <memberOf  key="model.phrase"/>
  </classes>
  <remarks>
    <p>The principles on which segmentation is carried out, and
        any special codes or attribute values used, should be defined
explicitly
        in the <gi>segmentation</gi> element of the <gi>encodingDesc</gi>
within
        the associated TEI header.</p>
  </remarks>
</classSpec>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identSynonyms"/>
   <classRef key="model.descLike"/>
  </alternate>
  <elementRef key="classes" minOccurs="0"/>
```

```
  <elementRef key="constraintSpec"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="attList" minOccurs="0"/>
  <elementRef key="exemplum" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="remarks" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="listRef" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element classSpec
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.identified.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   att.typed.attribute.subtype,
   attribute type { "model" | "atts" },
   attribute generate
   {
      list
      {
         (
            "alternation"
          | "sequence"
          | "sequenceOptional"
          | "sequenceOptionalRepeatable"
          | "sequenceRepeatable"
         ),
         (
            "alternation"
          | "sequence"
          | "sequenceOptional"
          | "sequenceOptionalRepeatable"
          | "sequenceRepeatable"
         )?,
         (
            "alternation"
          | "sequence"
          | "sequenceOptional"
          | "sequenceOptionalRepeatable"
          | "sequenceRepeatable"
         )?,
         (
            "alternation"
          | "sequence"
          | "sequenceOptional"
```

```
            | "sequenceOptionalRepeatable"
            | "sequenceRepeatable"
          )?,
          (
             "alternation"
            | "sequence"
            | "sequenceOptional"
            | "sequenceOptionalRepeatable"
            | "sequenceRepeatable"
          )?
       }
    }?,
    (
       ( model.identSynonyms | model.descLike )*,
       classes?,
       constraintSpec*,
       attList?,
       exemplum*,
       remarks*,
       listRef*
    )
}
```

---

## <climate>

(climate) contains information about the physical climate of a place. [*14.3.4.3. States, Traits, and Events*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*

  - att.canonical
    * *@key*

    * *@ref*

  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.placeStateLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc climate country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns place placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*certainty:* precision

*core:* bibl biblStruct desc head label listBibl note noteGrp p

*header:* biblFull

*linking:* ab

*msdescription:* msDesc

*namesdates:* climate

*Example*

```
<place xml:id="ROMA">
  <placeName>Rome</placeName>
<!-- ... -->
```

```
      <climate>
        <ab>
          <table>
            <head>24-hr Average Temperature</head>
            <row>
              <cell/>
              <cell  role="label">Jan</cell>
              <cell  role="label">Jun</cell>
              <cell  role="label">Dec</cell>
            </row>
            <row>
              <cell  role="label">°C</cell>
              <cell  role="data">7.1</cell>
              <cell  role="data">21.7</cell>
              <cell  role="data">8.3</cell>
            </row>
            <row>
              <cell  role="label">°F</cell>
              <cell  role="data">44.8</cell>
              <cell  role="data">71.1</cell>
              <cell  role="data">46.9</cell>
            </row>
          </table>
        </ab>
        <note>Taken from <bibl>
            <abbr>GHCN 2 Beta</abbr>: The Global Historical Climatology Network,
                version 2 beta, 1904 months between 1811 and 1980.
      <ptr  target="http://www.worldclimate.com/cgi-bin/data.pl?ref=N41E012+1202+0004058G2"/>
          </bibl>
        </note>
      </climate>
    </place>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="precision" minOccurs="0"
   maxOccurs="unbounded"/>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate>
   <classRef key="model.pLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <classRef key="model.labelLike"
    minOccurs="1" maxOccurs="unbounded"/>
  </alternate>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.noteLike"/>
   <classRef key="model.biblLike"/>
  </alternate>
  <elementRef key="climate" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element climate
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   (
      precision*,
      model.headLike*,
      ( model.pLike+ | model.labelLike+ ),
      ( model.noteLike | model.biblLike )*,
      climate*
   )
}
```

---

**<closer>** (closer) groups together salutations, datelines, bylines, and similar phrases appearing as a final group at the end of a division, especially of a letter. [*4.2.2. Openers and Closers 4.2. Elements Common to All Divisions*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*

    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.written
  - *@hand*

*Member of* model.divBottomPart

*Contained by*

*cmc:* post

*core:* lg list

*drama:* epilogue performance prologue

*figures:* figure table

*textstructure:* back body div div1 div2 div3 div4 div5 div6 div7 front group postscript

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q ref reg rs ruby sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* byline dateline salute signed

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<div type="letter">
  <p> perhaps you will favour me with a sight of it when convenient.</p>
  <closer>
    <salute>I remain, &c. &c.</salute>
    <signed>H. Colburn</signed>
  </closer>
</div>
```

*Example*

```
<div type="chapter">
  <p>
<!-- ... --> and his heart was going like mad and yes I said yes I will
Yes.</p>
  <closer>
    <dateline>
      <name type="place">Trieste-Zürich-Paris,</name>
      <date>1914–1921</date>
    </dateline>
  </closer>
</div>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <elementRef key="byline"/>
  <elementRef key="signed"/>
  <elementRef key="dateline"/>
  <elementRef key="salute"/>
  <classRef key="model.phrase"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element closer
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.cmc.attributes,
   att.written.attributes,
   (
      text
    | model.gLike    | byline    | signed    | dateline    | salute    | model.phrase    |
}
```

---

**<code>** contains literal code from some formal language such as a programming language. [*23.1.1. Phrase Level Terms*]

*Module* tagdocs — Documentation Elements

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@lang` (formal language) a name identifying the formal language in which the code is expressed.

*Status* Optional
*Datatype* teidata.word

*Member of* model.emphLike
*Contained by*
*analysis:* cl phr s span
*cmc:* post
*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit
*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castItem role roleDesc sound tech view
*figures:* cell figDesc
*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage
*iso-fs:* fDescr fsDescr
*linking:* ab seg
*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark
*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname
*spoken:* u writing
*tagdocs:* eg valDesc
*textcrit:* lem rdg wit witDetail witness
*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer
*transcr:* damage fw metamark mod restore retrace secl supplied surplus
*verse:* metSym rhyme
*May contain* Character data only
*Example*

```
<code  lang="JAVA"> Size fCheckbox1Size = new Size();
fCheckbox1Size.Height = 500;
fCheckbox1Size.Width = 500;
xCheckbox1.setSize(fCheckbox1Size);
</code>
```

*Content model* `<content> <textNode/></content>`
*Schema Declaration*

```
element code
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
```

```
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute lang { text }?,
   text
}
```

---

**<collation>** (collation) contains a description of how the leaves, bifolia, or similar objects are physically arranged. [*11.7.1. Object Description*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by:* *msdescription:* supportDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<collation>The written leaves preceded by an original flyleaf,
conjoint with the pastedown.</collation>
```

*Example*

```
<collation>
  <p>
    <formula>1-5.8 6.6 (catchword, f. 46, does not match following text)
        7-8.8 9.10, 11.2 (through f. 82) 12-14.8 15.8(-7)</formula>
    <catchwords>Catchwords are written horizontally in center
        or towards the right lower margin in various manners:
        in red ink for quires 1-6 (which are also signed in red
        ink with letters of the alphabet and arabic numerals);
        quires 7-9 in ink of text within yellow decorated frames;
        quire 10 in red decorated frame; quire 12 in ink of text;
        quire 13 with red decorative slashes; quire 14 added in
        cursive hand.</catchwords>
```

```
      </p>
    </collation>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element collation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.specialPara}
```

---

**<collection>** (collection) contains the name of a collection of manuscripts or other objects, not necessarily located within a single repository. [*11.4. The Manuscript Identifier*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*

    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
    * *@key*
    * *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Contained by msdescription:* altIdentifier msIdentifier
*namesdates:* objectIdentifier
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit
*dictionaries:* lang
*figures:* figure notatedMusic
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef tag val
*textcrit:* app witDetail
*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data
*Example*

```
<msIdentifier>
  <country>USA</country>
  <region>California</region>
  <settlement>San Marino</settlement>
  <repository>Huntington Library</repository>
  <collection>Ellesmere</collection>
  <idno>El 26 C 9</idno>
  <msName>The Ellesmere Chaucer</msName>
</msIdentifier>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element collection
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq.limited}
```

---

**<colloc>** (collocate) contains any sequence of words that co-occur with the headword with significant frequency. [*10.3.2. Grammatical Information*]

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*

  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    - *@datcat*
    - *@valueDatcat*
    - *@targetDatcat*
  - att.lexicographic.normalized
    - *@norm*
    - *@orig*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.entryPart model.lexicalRefinement
*Contained by*
*core:* cit
*dictionaries:* dictScrap entryFree form gramGrp
*namesdates:* nym
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<entry>
  <form>
    <orth>médire</orth>
  </form>
  <gramGrp>
    <colloc>de</colloc>
  </gramGrp>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element colloc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
```

```
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.typed.attributes,
   macro.paraContent}
```

---

**<colophon>** (colophon) contains the *colophon* of an item: that is, a statement providing information regarding the date, place, agency, or reason for production of the manuscript or other object. [*11.6.1. The msItem and msItemStruct Elements*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

    * *@source*

  - att.msExcerpt
    - *@defective*

*Member of* model.msQuoteLike

*Contained by*

*msdescription:* msItem msItemStruct

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<colophon>Ricardus Franciscus Scripsit Anno Domini
1447.</colophon>
```

*Example*

```
<colophon>Explicit expliceat/scriptor ludere eat.</colophon>
```

*Example*

```
<colophon>Explicit venenum viciorum domini illius, qui comparavit Anno
domini Millessimo Trecentesimo nonagesimo primo, Sabbato in festo
sancte Marthe virginis gloriose. Laus tibi criste quia finitur
libellus iste.</colophon>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
```

```
</content>
```

*Schema Declaration*

```
element colophon
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.msExcerpt.attributes,
   macro.phraseSeq}
```

---

**<cond>** (conditional feature-structure constraint) defines a conditional feature-structure constraint; the consequent and the antecedent are specified as feature structures or feature-structure collections; the constraint is satisfied if both the antecedent and the consequent subsume a given feature structure, or if the antecedent does not. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Contained by:* *iso-fs:* fsConstraints

*May contain*

*iso-fs:* f fs then

*Note* May contain an antecedent feature structure, an empty <then> element, and a consequent feature structure.

*Example*

```
<cond>
  <fs>
    <f  name="BAR">
      <symbol  value="1"/>
    </f>
  </fs>
  <then/>
  <fs>
    <f  name="SUBCAT">
      <binary  value="false"/>
    </f>
  </fs>
</cond>
```

*Content model*

```
<content>
 <sequence>
  <alternate>
   <elementRef key="fs"/>
   <elementRef key="f"/>
  </alternate>
  <elementRef key="then"/>
  <alternate>
   <elementRef key="fs"/>
   <elementRef key="f"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element cond
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
```

```
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( ( fs | f ), then, ( fs | f ) )
}
```

---

## <condition>

(condition) contains a description of the physical condition of the manuscript or object. [*11.7.1.5. Condition*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by:* *msdescription:* binding bindingDesc sealDesc supportDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<condition>
  <p>There are lacunae in three places in this
     manuscript. After 14v two
     leaves has been cut out and narrow strips leaves remains in the spine.
After
     68v one gathering is missing and after 101v at least one gathering of 8
leaves
     has been lost. </p>
  <p>Several leaves are damaged with tears or holes or have a
     irregular shape. Some of the damages do not allow the lines to be of full
     length and they are apparently older than the script. There are tears on
fol.
     2r-v, 9r-v, 10r-v, 15r-18v, 19r-v, 20r-22v, 23r-v, 24r-28v, 30r-v,
32r-35v,
     37r-v, 38r-v, 40r-43v, 45r-47v, 49r-v, 51r-v, 53r-60v, 67r-v, 68r-v,
70r-v,
     74r-80v, 82r-v, 86r-v, 88r-v, 89r-v, 95r-v, 97r-98v 99r-v, 100r-v. On fol.
98
```

```
      the corner has been torn off. Several leaves are in a bad condition due to
      moist and wear, and have become dark, bleached or
      wrinkled. </p>
    <p>The script has been
      touched up in the 17th century with black ink. The touching up on the
  following
      fols. was done by
    <name>Bishop Brynjólf Sveinsson</name>: 1v, 3r, 4r, 5r,
      6v, 8v,9r, 10r, 14r, 14v, 22r,30v, 36r-52v, 72v, 77r,78r,103r, 104r,. An
      AM-note says according to the lawman
    <name>Sigurður Björnsson</name> that the rest of the
      touching up was done by himself and another lawman
    <name>Sigurður Jónsson</name>.
    <name>Sigurður Björnsson</name> did the touching up
      on the following fols.: 46v, 47r, 48r, 49r-v, 50r, 52r-v.
    <name>Sigurður Jónsson</name> did the rest of the
      touching up in the section 36r-59r containing
    <title>Bretasögur</title>
    </p>
  </condition>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element condition
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.specialPara}
```

---

## <constitution>

<constitution> (constitution) describes the internal composition of a text or text sample, for example as fragmentary, complete, etc. [*16.2.1. The Text Description*]

*Module* corpus — Language Corpora

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

• att.typed
- ~~type~~
- @subtype

`@type` specifies how the text was constituted.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **single** a single complete text[Default]

**composite** a text made by combining several smaller items, each individually complete

**frags** (fragments) a text made by combining several smaller, not necessarily complete, items

**unknown** composition unknown or unspecified

*Member of* model.textDescPart

*Contained by*

*corpus:* textDesc

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data

*Note* The function of this element seems to overlap with both the *org* attribute on  and the <samplingDecl> in the <encodingDesc>.

*Example*

```
<constitution type="frags">Prologues only.</constitution>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element constitution
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { "single" | "composite" | "frags" | "unknown" }?,
   macro.phraseSeq.limited}
```

---

## <constraint> (constraint rules) the formal rules of a constraint [*23.5. Element Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Contained by* tei: constraintSpec
*May contain* ANY
*Example*

```
<constraintSpec ident="isoconstraint"
  scheme="schematron"
   xmlns:sch="http://purl.oclc.org/dsdl/schematron">
  <constraint>
    <sch:rule context="tei:constraint">
      <sch:assert test="tei:fileDesc/tei:titleStmt/tei:title[ @type eq
'main']">a main title must be supplied</sch:assert>
    </sch:rule>
  </constraint>
</constraintSpec>
```

*Example*

```
<constraintSpec ident="local"
  scheme="private">
```

```
  <constraint>
     (output = (ident(title) ident(author) "You must supply a title or an
author"))
  </constraint>
</constraintSpec>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <anyElement/>
 </alternate>
</content>
```

*Schema Declaration*

```
element constraint
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( text | anyElement_constraint_5 )*
}
```

---

**<constraintDecl>** (constraint declaration) contains declarations pertaining to formal constraints expressed elsewhere in <constraintSpec> elements [*23.5.2. Additional Constraints*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@scheme` supplies the name of the language to which the declarations herein apply
*Status* Required
*Datatype* teidata.enumerated
*Suggested values include:* **schematron** (ISO Schematron)
*Note* The declarations contained in a particular <constraintDecl> (including *queryBinding*, if present) apply to the <constraintSpec> elements whose *scheme* matches the *scheme* of the <constraintDecl>.

`@queryBinding` (query language binding) specifies the query language binding for rule-based schema expressions in <constraintSpec> elements that have a matching *scheme* attribute
*Status* Recommended
*Datatype* teidata.enumerated
*Suggested values include:* **exslt**
**stx**
**xslt**
**xslt2**
**xslt3**
**xpath**
**xpath2**
**xpath3**
**xpath31**
**xquery**
**xquery3**
**xquery31**
*Note* The suggested values above are the values reserved by the Schematron specification. Only exslt, stx, xslt, xslt2, xslt3, xpath2, and xpath3 are defined by the specification. Most processors only support one or more of xslt, xslt2, and xslt3.

*Member of* model.encodingDescPart
*Contained by*

*header:* encodingDesc
*tagdocs:* schemaSpec
*May contain* ANY
*Example*

```
<constraintDecl scheme="schematron"
  queryBinding="xslt3"
   xmlns:sch="http://purl.oclc.org/dsdl/schematron">
  <sch:ns prefix="wwp"
    uri="http://www.wwp.northeastern.edu/ns/textbase"/>
</constraintDecl>
```

*Schematron* <sch:rule context=”/*[ count( //tei:constraintDecl ) gt 1 ]”> <sch:let name=”schemes” value=”//tei:constraintDecl[ not(ancestor::*[local-name(.) eq ’egXML’] ) ]/@scheme”/> <sch:let name=”disctinct_schemes” value=”distinct-values( $schemes )”/> <sch:assert test=”count( $schemes ) eq count( $disctinct_schemes )”> Each <constraintDecl> element should have a @scheme attribute that is distinct from that of all the other <constraintDecl>s. </sch:assert> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identEquiv"/>
   <classRef key="model.descLike"/>
  </alternate>
  <anyElement minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element constraintDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute scheme { "schematron" },
   attribute queryBinding
   {
      "exslt"
    | "stx"
    | "xslt"
    | "xslt2"
    | "xslt3"
    | "xpath"
    | "xpath2"
    | "xpath3"
    | "xpath31"
    | "xquery"
    | "xquery3"
```

```
    | "xquery31"
   }?,
   ( ( model.identEquiv | model.descLike )*, anyElement_constraintDecl_4* )
}
```

---

**<constraintSpec>** (constraint on schema) contains a formal constraint, typically expressed in a rule-based schema language, to which a construct must conform in order to be considered valid [*23.5. Element Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.identified
  - *@ident*
  - *@predeclare*

– *@module*

– att.combinable
* *@mode*

* att.deprecated
· *@validUntil*

• att.typed
– ~~type~~
– @subtype

`@type` characterizes the <constraintSpec> element in some sense; used to indicate when a <constraintSpec> warns about a deprecated construct.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **deprecationWarning** Indicates that this constraint specification warns that some other construct in the schema is deprecated.

`@scheme` supplies the name of the language in which the constraints are defined
*Status* Optional
*Datatype* teidata.enumerated
*Schematron* <sch:rule context="tei:constraintSpec[ @mode = ('add','replace') or not( @mode ) ]"> <sch:assert test="@scheme">The @scheme attribute of <constraintSpec> is required when the @mode is <sch:value-of select="if (@mode) then concat('"',@mode,'"') else 'not specified'"/> (here on "<sch:value-of select="@ident"/>")</sch:assert> </sch:rule>
*Suggested values include:* **schematron** (ISO Schematron)
*Note* These Guidelines no longer recommend, and the current TEI stylesheets no longer support, the use of Schematron 1.x within <constraintSpec>. Thus the value schematron is used to indicate that ISO Schematron is used within the <constraintSpec>.

The *scheme* attribute is required when the value of *mode* is add or replace. The *scheme* attribute is permitted when the value of *mode* is delete, but these Guidelines make no reccomendation for what a processor should do if its value does not match that of the <constraintSpec> being deleted.

*Member of* model.oddDecl
*Contained by*
*cmc:* post
*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr
*drama:* camera caption castList epilogue performance prologue set sound tech view
*figures:* cell figDesc figure
*header:* change handNote licence rendition scriptNote tagUsage
*iso-fs:* fDescr fsDescr
*linking:* ab seg
*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote
*namesdates:* occupation

*spoken:* u writing

*tagdocs:* attDef classSpec dataSpec elementSpec macroSpec schemaSpec specGrp

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc gloss

*tagdocs:* altIdent constraint equiv

*Note* A child <constraint> is required when the *mode* attribute has the value replace or add (or is not specified, as add is the default). No child elements are permitted when the *mode* attribute has the value delete. At least one child element is required when the *mode* attribute has the value change.

*Example* This constraint uses Schematron to enforce the presence of the *spanTo* attribute (which comes from an attribute class) on the <addSpan> element:

```
<constraintSpec ident="demo-c1"
  scheme="schematron"
   xmlns:sch="http://purl.oclc.org/dsdl/schematron">
  <desc>Enforce the presence of the <att>spanTo</att> attribute</desc>
  <constraint>
    <sch:rule context="tei:addSpan">
      <sch:assert test="@spanTo">The spanTo= attribute of <sch:name/> is
required.</sch:assert>
    </sch:rule>
  </constraint>
</constraintSpec>
```

*Example*

```
<constraintSpec ident="demo-c2"
  scheme="schematron"
   xmlns:sch="http://purl.oclc.org/dsdl/schematron">
  <desc>Implement an accessibility rule which says that pictures
     should have textual explanations</desc>
  <constraint>
    <sch:rule context="tei:figure">
      <sch:report test="not( tei:figDesc | tei:head )">You should
            provide information in a figure from which
            we can construct an alt attribute in HTML</sch:report>
    </sch:rule>
  </constraint>
</constraintSpec>
```

*Example* This constraint uses SPITBOL (a language which is not expressed in XML) to check whether the title and author are identical:

```
<constraintSpec ident="author_ne_title"
  scheme="SPITBOL">
  <constraint>
     (output = leq(title,author) "title and author cannot be the same")
  </constraint>
</constraintSpec>
```

*Schematron* <sch:rule context="tei:constraintSpec[ @mode eq 'delete']"> <sch:report test="child::*">This

constraintSpec element has a mode= of ”delete” even though it has child elements. Change the mode= to ”add”, ”change”, or ”replace”, or remove the child elements.</sch:report> </sch:rule>
<sch:rule context=”tei:constraintSpec[ @mode eq ’change’]”> <sch:assert test=”child::*”>This constraintSpec element has a mode= of ”change”, but does not have any child elements. Specify child elements, or change the mode= to ”delete”.</sch:assert> </sch:rule>
<sch:rule context=”tei:constraintSpec[ @mode = (’add’,’replace’) ]”>
<sch:assert test=”child::tei:constraint”>This constraintSpec element has a mode= of ”<sch:value-of select=”@mode”/>”, but does not have a child ’constraint’ element. Use a child ’constraint’ element or change the mode= to ”delete” or ”change”.</sch:assert> </sch:rule>

*Schematron* Relationship between scheme attribute and contents: Schematron 1.x
<sch:rule context=”tei:constraintSpec”> <sch:report test=”tei:constraint/sch1x:* and @scheme = (’isoschematron’,’schematron’)”>Rules in the Schematron 1.* language must be inside a constraintSpec with a value other than ’schematron’ or ’isoschematron’ on the scheme attribute</sch:report> </sch:rule>

*Schematron* Relationship between scheme attribute and contents: ISO Schematron
<sch:rule context=”tei:constraintSpec[ @mode = (’add’,’replace’) or not( @mode ) ]”>
<sch:report test=”tei:constraint/sch:* and not( @scheme eq ’schematron’)”>Rules in the ISO Schematron language must be inside a constraintSpec with the value ’schematron’ on the scheme attribute</sch:report> </sch:rule>

*Schematron* <sch:rule context=”(tei:constraintSpec|tei:constraintDecl)[ @scheme eq ’schematron’][ .//sch:assert | .//sch:report ]”> <sch:let name=”assertsHaveContext” value=”for $a in .//sch:assert return exists( $a/ancestor::sch:rule/@context )”/> <sch:let name=”reportsHaveContext” value=”for $r in .//sch:report return exists( $r/ancestor::sch:rule/@context )”/> <sch:report test=”( $assertsHaveContext, $reportsHaveContext ) = false()”>An <sch:assert> or <sch:report> element must be a descendant of an <sch:rule> that has a @context; one or more inside this <sch:value-of select=”local-name(.)”/> are not.</sch:report> </sch:rule>

*Schematron* <sch:rule context=”tei:constraintSpec[ @mode eq ’add’ or not( @mode ) ]”>
<sch:let name=”myIdent” value=”normalize-space(@ident)”/>
<sch:report test=”preceding::tei:constraintSpec[ normalize-space(@ident) eq $myIdent ]”> The @ident of ’constraintSpec’ should be unique; this one (<sch:value-of select=”$myIdent”/>) is the same as that of a previous ’constraintSpec’. </sch:report> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identSynonyms"/>
   <classRef key="model.descLike"/>
  </alternate>
  <elementRef key="constraint"
   minOccurs="0" maxOccurs="1"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element constraintSpec
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
```

```
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.identified.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   att.typed.attribute.subtype,
   attribute type { "deprecationWarning" }?,
   attribute scheme { "schematron" }?,
   ( ( model.identSynonyms | model.descLike )*, constraint? )
}
```

---

**<content>** (content model) contains a declaration of the intended content model for the element (or other construct) being specified. [*23.5. Element Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

* *@source*

`@autoPrefix` controls whether or not pattern names generated in the corresponding RELAX NG schema source are automatically prefixed to avoid potential nameclashes.
*Status* Optional
*Datatype* teidata.truthValue
*Legal values are:* **true** Each name referenced in e.g. an `<rng:ref>` element within a content model is automatically prefixed by the value of the *prefix* attribute on the current <schemaSpec>[Default]
**false** No prefixes are added: any prefix required by the value of the *prefix* attribute on the current <schemaSpec> must therefore be supplied explicitly, as appropriate.

*Contained by:* tagdocs: dataSpec elementSpec macroSpec moduleRef
*May contain* ANY
*Note* It is required that the <content> element has only one child element. If several RELAX NG elements are desired, they must be wrapped in a `<rng:div>`.
*Example* This sample <content> element indicates that the element being specified has no content:

```
<content>
  <empty/>
</content>
```

*Example* This <content> element defines a content model, expressed directly in the TEI ODD language, that allows either a sequence of paragraphs or a series of <msItem> elements optionally preceded by a <summary>:

```
<content>
  <alternate>
    <classRef key="model.pLike"
      maxOccurs="unbounded"/>
    <sequence>
      <elementRef key="summary" minOccurs="0"
        maxOccurs="1"/>
      <elementRef key="msItem"
        maxOccurs="unbounded"/>
    </sequence>
  </alternate>
</content>
```

*Example* This sample <content> element defines a content model, expressed in the RELAX NG schema language, that allows either a sequence of paragraphs or a series of <msItem> elements optionally preceded by a <summary>:

```
<content>
  <rng:choice>
    <rng:oneOrMore>
      <rng:ref name="model.pLike"/>
    </rng:oneOrMore>
    <rng:group>
      <rng:optional>
        <rng:ref name="summary"/>
      </rng:optional>
      <rng:oneOrMore>
        <rng:ref name="msItem"/>
      </rng:oneOrMore>
    </rng:group>
```

```
      </rng:choice>
    </content>
```

*Content model*

```
<content>
 <alternate>
  <elementRef key="valList" minOccurs="1"
   maxOccurs="1"/>
  <classRef key="model.contentPart"
   minOccurs="1" maxOccurs="1"/>
  <anyElement minOccurs="1" maxOccurs="1"
   require="http://relaxng.org/ns/compatibility/annotations/1.0 http://relaxng.org/ns/stru
 </alternate>
</content>
```

*Schema Declaration*

```
element content
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute autoPrefix { "true" | "false" }?,
   ( valList | model.contentPart | anyElement_content_3 )
}
```

---

**<conversion>** defines how to calculate one unit of measure in terms of another. [*2.3.9. The Unit Declaration*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.formula
  - *@formula*
- att.locatable

– *@where*

`@fromUnit` indicates a source unit of measure that is to be converted into another unit indicated in *toUnit*.
*Status* Required
*Datatype* teidata.pointer

`@toUnit` the target unit of measurement for a conversion from a source unit referenced in *fromUnit*.
*Status* Required
*Datatype* teidata.pointer

*Contained by:* unitDef

*May contain* Empty element

*Note* The conversion element is designed to store information about converting from one unit of measurement to another. The *formula* attribute holds an XPath expression that indicates how the measurement system in *fromUnit* is converted to the system in *toUnit*. Do not confuse the usage of the dating attributes (*from* and *to*) in the examples with the attributes (*fromUnit* and *toUnit*) designed to reference units of measure.

*Example*

```
<conversion fromUnit="#shilling"
  toUnit="#pence" formula="$fromUnit * 12" from="1707"
  to="1971"/>
```

*Example*

```
<conversion fromUnit="#pound"
  toUnit="#shilling" formula="$fromUnit * 20" from="1701"
  to="1971"/>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element conversion
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.formula.attributes,
   att.locatable.attributes,
   attribute fromUnit { text },
   attribute toUnit { text },
   empty
}
```

**<corr>** (correction) contains the correct form of a passage apparently erroneous in the copy text. [*3.5.1. Apparent Errors*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.editLike
  - *@evidence*
  - *@instant*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.choicePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example* If all that is desired is to call attention to the fact that the copy text has been corrected, <corr> may be used alone:

```
I don't know,
Juan. It's so far in the past now — how <corr>can we</corr> prove
or disprove anyone's theories?
```

*Example* It is also possible, using the <choice> and <sic> elements, to provide an uncorrected reading:

```
I don't know, Juan. It's so far in the past now —
how <choice>
  <sic>we can</sic>
  <corr>can we</corr>
</choice> prove or
disprove anyone's theories?
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element corr
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.editLike.attributes,
   att.typed.attributes,
   macro.paraContent}
```

---

## <correction>

**<correction>** (correction principles) states how and under what circumstances corrections have been made in the text. [*2.3.3. The Editorial Practices Declaration 16.3.2. Declarable Elements*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.declarable
  - *@default*

`@status` indicates the degree of correction applied to the text.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **high** the text has been thoroughly checked and proofread.

**medium** the text has been checked at least once.

**low** the text has not been checked.

**unknown** the correction status of the text is unknown.

`@method` indicates the method adopted to indicate corrections within the text.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **silent** corrections have been made silently[Default]

**markup** corrections have been represented using markup

*Member of* model.editorialDeclPart

*Contained by*

*header:* editorialDecl

*May contain*

*core:* p

*linking:* ab

*Note* May be used to note the results of proof reading the text against its original, indicating (for example) whether discrepancies have been silently rectified, or recorded using the editorial tags described in section *3.5. Simple Editorial Changes*.

*Example*

```
<correction>
  <p>Errors in transcription controlled by using the WordPerfect spelling
checker, with a user
    defined dictionary of 500 extra words taken from Chambers Twentieth
Century
    Dictionary.</p>
</correction>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:correction"/> </sch:pattern>

*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element correction
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   attribute status { "high" | "medium" | "low" | "unknown" }?,
   attribute method { "silent" | "markup" }?,
   model.pLike+
}
```

---

## <correspAction>

<correspAction> (correspondence action) contains a structured description of the place, the name of a person/organization and the date related to the sending/receiving of a message or any other action related to the correspondence. [*2.4.6. Correspondence Description*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.sortable
  - *@sortKey*
- att.typed
  - ~~type~~
  - @subtype

`@type` describes the nature of the action.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **sent** information concerning the sending or dispatch of a message.

**received** information concerning the receipt of a message.

**transmitted** information concerning the transmission of a message, i.e. between the dispatch and the next receipt, redirect or forwarding.

**redirected** information concerning the redirection of an unread message.

**forwarded** information concerning the forwarding of a message.

*Member of* model.correspDescPart

*Contained by*

*header:* correspDesc

*May contain*

*core:* address date email name note noteGrp p rs time

*dictionaries:* lang

*header:* idno

*linking:* ab

*namesdates:* addName affiliation bloc climate country district eventName forename genName geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*Example*

```
<correspAction type="sent">
  <persName>Adelbert von Chamisso</persName>
  <settlement>Vertus</settlement>
  <date when="1807-01-29"/>
</correspAction>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.correspActionPart"
   minOccurs="1" maxOccurs="unbounded"/>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element correspAction
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.sortable.attributes,
   att.typed.attribute.subtype,
   attribute type
   {
      "sent" | "received" | "transmitted" | "redirected" | "forwarded"
   }?,
   ( model.correspActionPart+ | model.pLike+ )
}
```

---

## <correspContext>

<correspContext> (correspondence context) provides references to preceding or following correspondence related to this piece of correspondence. [*2.4.6. Correspondence Description*]

*Module* header — The TEI Header

*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  * *@ana*
- – att.global.change
  * *@change*
- – att.global.facs
  * *@facs*
- – att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- – att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- – att.global.responsibility
  * *@cert*
  * *@resp*
- – att.global.source
  * *@source*

*Member of* model.correspDescPart
*Contained by*
*header:* correspDesc
*May contain*
*core:* note noteGrp p ptr ref
*linking:* ab
*tagdocs:* listRef
*Example*

```
<correspContext>
  <ptr type="next" subtype="toAuthor"
    target="http://tei.ibi.hu-berlin.de/berliner-intellektuelle/manuscript?Brief101Varnhagenar
  <ptr type="prev" subtype="fromAuthor"
```

```
      target="http://tei.ibi.hu-berlin.de/berliner-intellektuelle/manuscript?Brief103BoeckhanVar
  </correspContext>
```

*Example*

```
  <correspContext>
    <ref  type="prev"
      target="http://weber-gesamtausgabe.de/A040962"> Previous letter of
    <persName>Carl Maria von Weber</persName> to
    <persName>Caroline Brandt</persName>:
    <date  when="1816-12-30">December 30, 1816</date>
    </ref>
    <ref  type="next"
      target="http://weber-gesamtausgabe.de/A041003"> Next letter of
    <persName>Carl Maria von Weber</persName> to
    <persName>Caroline Brandt</persName>:
    <date  when="1817-01-05">January 5, 1817</date>
    </ref>
  </correspContext>
```

*Content model*

```
<content>
 <classRef key="model.correspContextPart"
  minOccurs="1" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element correspContext
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   model.correspContextPart+
}
```

---

## <correspDesc>

**<correspDesc>** (correspondence description) contains a description of the actions related to one act of correspondence. [*2.4.6. Correspondence Description*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.declarable
  - *@default*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.profileDescPart
*Contained by*
*header:* profileDesc
*May contain*
*core:* note noteGrp p
*header:* correspAction correspContext
*linking:* ab
*Example*

```
<correspDesc>
  <correspAction type="sent">
```

```
      <persName>Carl Maria von Weber</persName>
      <settlement>Dresden</settlement>
      <date  when="1817-06-23">23 June 1817</date>
    </correspAction>
    <correspAction  type="received">
      <persName>Caroline Brandt</persName>
      <settlement>Prag</settlement>
    </correspAction>
    <correspContext>
      <ref  type="prev"
        target="http://www.weber-gesamtausgabe.de/A041209">Previous letter of
      <persName>Carl Maria von Weber</persName>
          to <persName>Caroline Brandt</persName>:
      <date  from="1817-06-19"  to="1817-06-20">June 19/20, 1817</date>
      </ref>
      <ref  type="next"
        target="http://www.weber-gesamtausgabe.de/A041217">Next letter of
      <persName>Carl Maria von Weber</persName> to
      <persName>Caroline Brandt</persName>:
      <date  when="1817-06-27">June 27, 1817</date>
      </ref>
    </correspContext>
  </correspDesc>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:correspDesc"/> </sch:pattern>

*Content model*

```
<content>
 <alternate>
  <classRef key="model.correspDescPart"
   minOccurs="1" maxOccurs="unbounded"/>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element correspDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.declarable.attributes,
   att.typed.attributes,
   ( model.correspDescPart+ | model.pLike+ )
}
```

**<country>** (country) contains the name of a geo-political unit, such as a nation, country, colony, or commonwealth, larger than or administratively superior to a region and smaller than a bloc. [*14.2.3. Place Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*

    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
    * *@key*
    * *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.placeNamePart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage unitDef

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions altIdentifier catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material

msIdentifier musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectIdentifier objectName occupation offset org orgName persName persPronouns place placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The recommended source for codes to represent coded country names is ISO 3166.

*Example*

```
<country key="DK">Denmark</country>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element country
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<creation>** (creation) contains information about the creation of a text. [*2.4.1. Creation 2.4. The Profile Description*]

*Module* header — The TEI Header

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition

    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*

*Member of* model.profileDescPart

*Contained by*

*header:* profileDesc

*May contain*

*core:* abbr address choice date distinct email emph expan foreign gloss hi measure measureGrp mentioned name num ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*header:* idno listChange

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*tagdocs:* att code gi ident listRef tag val

*transcr:* am ex subst

character data

*Note* The <creation> element may be used to record details of a text's creation, e.g. the date and place it was composed, if these are of interest.

It may also contain a more structured account of the various stages or revisions associated with the evolution of a text; this should be encoded using the <listChange> element. It should not be confused with the <publicationStmt> element, which records date and place of publication.

*Example*

```
<creation>
  <date>Before 1987</date>
</creation>
```

*Example*

```
<creation>
  <date  when="1988-07-10">10 July 1988</date>
</creation>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.limitedPhrase"/>
  <elementRef key="listChange"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element creation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   ( text | model.limitedPhrase | listChange )*
}
```

---

**<cRefPattern>** (canonical reference pattern) specifies an expression and replacement pattern for transforming a canonical reference into a URI. [*2.3.6.3. Milestone Method 2.3.6. The Reference System Declaration 2.3.6.2. Search-and-Replace Method*]

*Module* header — The TEI Header
*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.patternReplacement
  - *@matchPattern*
  - *@replacementPattern*

*Contained by:* refsDecl
*May contain*
*core:* p
*linking:* ab
*Note* The result of the substitution may be either an absolute or a relative URI reference. In the latter case it is combined with the value of *xml:base* in force at the place where the *cRef* attribute occurs to form an absolute URI in the usual manner as prescribed by XML Base.
*Example*

```
<cRefPattern matchPattern="([1-9A-Za-z]+)\s+([0-9]+):([0-9]+)"
  replacementPattern="#xpath(//div[@type='book'][@n='$1']/div[@type='chap'][@n='$2']/div[@type
```

*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="0"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element cRefPattern
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.patternReplacement.attributes,
   model.pLike*
}
```

---

## <custEvent>

(custodial event) describes a single event during the custodial history of a manuscript or other object. [*11.9.1.2. Availability and Custodial History*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*

    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.typed
  - *@type*
  - *@subtype*

*Contained by:* *model.descriptionLike:* custodialHist
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<custEvent  type="photography">Photographed by David Cooper on <date>12 Dec
1964</date>
</custEvent>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element custEvent
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
```

```
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.typed.attributes,
   macro.specialPara}
```

---

**<custodialHist>** (custodial history) contains a description of a manuscript or other object's custodial history, either as running prose or as a series of dated custodial events. [*11.9.1.2. Availability and Custodial History*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by:* *msdescription:* adminInfo
*May contain*
*core:* p
*linking:* ab
*msdescription:* custEvent
*Example*

```
<custodialHist>
  <custEvent  type="conservation"
    notBefore="1961-03"  notAfter="1963-02">Conserved between March 1961 and
February 1963 at
     Birgitte Dalls Konserveringsværksted.</custEvent>
  <custEvent  type="photography"
    notBefore="1988-05-01"  notAfter="1988-05-30">Photographed in
     May 1988 by AMI/FA.</custEvent>
  <custEvent  type="transfer-dispatch"
    notBefore="1989-11-13"  notAfter="1989-11-13">Dispatched to Iceland
     13 November 1989.</custEvent>
</custodialHist>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <elementRef key="custEvent" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element custodialHist
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.pLike+ | custEvent+ )
}
```

---

## <damage>

(damage) contains an area of damage to the text witness. [*12.3.3.1. Damage, Illegibility, and Supplied Text*]

*Module* transcr — Representation of Primary Sources
*Attributes*
- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.damaged
  - *@agent*
  - *@degree*
  - *@group*

- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    - *@atLeast*
    - *@atMost*

- *@min*
- *@max*
- *@confidence*

– att.written
* *@hand*

• att.typed
– *@type*
– *@subtype*

*Member of* model.linePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Since damage to text witnesses frequently makes them harder to read, the <damage> element will often contain an <unclear> element. If the damaged area is not continuous (e.g. a stain affecting several strings of text), the *group* attribute may be used to group together several related <damage> elements; alternatively the <join> element may be used to indicate which <damage> and <unclear> elements are part of the same physical phenomenon.

The <damage>, <gap>, <del>, <unclear> and <supplied> elements may be closely allied in use. See section *12.3.3.2. Use of the gap, del, damage, unclear, and supplied Elements in Combination* for discussion of which element is appropriate for which circumstance.

*Example*

```
<l>The Moving Finger wri<damage  agent="water"  group="1">es; and</damage>
having writ,</l>
<l>Moves <damage  agent="water"  group="1">
    <supplied>on: nor all your</supplied>
  </damage> Piety nor Wit</l>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element damage
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.damaged.attributes,
```

```
   att.dimensions.attributes,
   att.ranging.attributes,
   att.written.attributes,
   att.typed.attributes,
   macro.paraContent}
```

---

## <damageSpan>

(damaged span of text) marks the beginning of a longer sequence of text which is damaged in some way but still legible. [*12.3.3.1. Damage, Illegibility, and Supplied Text*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.damaged
  - *@agent*
  - *@degree*

- *@group*

- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

- att.ranging
  - *@atLeast*
  - *@atMost*
  - *@min*
  - *@max*
  - *@confidence*

- att.written
  - *@hand*

- att.spanning
  - *@spanTo*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.global.edit

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName

occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Note* Both the beginning and ending of the damaged sequence must be marked: the beginning by the <damageSpan> element, the ending by the target of the *spanTo* attribute: if no other element available, the <anchor> element may be used for this purpose.

The damaged text must be at least partially legible, in order for the encoder to be able to transcribe it. If it is not legible at all, the <damageSpan> element should not be used. Rather, the <gap> or <unclear> element should be employed, with the value of the *reason* attribute giving the cause. See further sections *12.3.3.1. Damage, Illegibility, and Supplied Text* and *12.3.3.2. Use of the gap, del, damage, unclear, and supplied Elements in Combination.*

*Example*

```
<p>Paragraph partially damaged. This is the undamaged
portion <damageSpan  spanTo="#a34"/>and this the damaged
portion of the paragraph.</p>
<p>This paragraph is entirely damaged.</p>
<p>Paragraph partially damaged; in the middle of this
paragraph the damage ends and the anchor point marks
the start of the <anchor  xml:id="a34"/> undamaged part of the text. ...</p>
```

*Schematron* <sch:rule context="tei:damageSpan"> <sch:assert test="@spanTo">The @spanTo attribute of <sch:name/> is required.</sch:assert> </sch:rule>

*Schematron* <sch:rule context="tei:damageSpan"> <sch:assert test="@spanTo">L'attribut spanTo est requis.</sch:assert> </sch:rule>

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element damageSpan
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.damaged.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.written.attributes,
   att.spanning.attributes,
   att.typed.attributes,
```

```
   empty
}
```

---

**<dataFacet>** Restricts the value of the strings used to represent values of a datatype, according to XML Schema Part 2: Datatypes Second Edition. [*23.5.3.1. Datatypes*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

`@name` the name of the facet.

*Status* Required

*Datatype* teidata.word

*Legal values are:* **length**
**minLength**

**maxLength**
**pattern**
**enumeration**
**whiteSpace**
**maxInclusive**
**minInclusive**
**maxExclusive**
**minExclusive**
**totalDigits**
**fractionDigits**

`@value` the facet value.
*Status* Required
*Datatype* string

*Contained by:* dataRef
*May contain* Empty element
*Note* This element is only allowed when the parent <dataRef> refers with *name* to a datatype from the specification XML Schema Part 2: Datatypes Second Edition.
*Example*

```
<datatype>
  <dataRef  name="decimal">
    <dataFacet  name="maxInclusive"
      value="360.0"/>
    <dataFacet  name="minInclusive"
      value="-360.0"/>
  </dataRef>
</datatype>
```

*Example*

```
<dataSpec  ident="data.integerExample">
  <desc>datatype used for attributes taking an integer value between 0 and
99</desc>
  <content>
    <dataRef  name="nonNegativeInteger">
      <dataFacet  name="maxInclusive"
        value="99"/>
      <dataFacet  name="minInclusive"  value="0"/>
    </dataRef>
  </content>
</dataSpec>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element dataFacet
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
```

```
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute name
   {
      "length"
    | "minLength"
    | "maxLength"
    | "pattern"
    | "enumeration"
    | "whiteSpace"
    | "maxInclusive"
    | "minInclusive"
    | "maxExclusive"
    | "minExclusive"
    | "totalDigits"
    | "fractionDigits"
   },
   attribute value { text },
   empty
}
```

---

**<dataRef>** identifies the datatype of an attribute value, either by referencing an item in an externally defined datatype library, or by pointing to a TEI-defined data specification [*23.5.3.1. Datatypes*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@key` the identifier used for this datatype specification
*Status* Optional
*Datatype* teidata.xmlName

`@name` the name of a datatype in the list provided by XML Schema Part 2: Datatypes Second Edition
*Status* Optional
*Datatype* teidata.xmlName

`@ref` a pointer to a datatype defined in some datatype library
*Status* Optional
*Datatype* teidata.pointer

`@restriction` supplies a string representing a regular expression providing additional constraints on the strings used to represent values of this datatype
*Status* Optional
*Datatype* teidata.pattern

*Member of* model.contentPart model.oddRef

*Contained by*

*tagdocs:* alternate content datatype schemaSpec sequence specGrp

*May contain*

*tagdocs:* dataFacet

*Note* Only one of the attributes *key*, *name*, and *ref* may be used on any given instance of <dataRef>.

Neither a *restriction* attribute nor a <dataFacet> child element may be used unless the <dataRef> refers to a datatype from the specification XML Schema Part 2: Datatypes Second Edition with a *name* attribute.

*Example*

```
<schemaSpec ident="myTEI"
 source="http://www.tei-c.org/Vault/P5/current/xml/tei/odd/p5subset.xml">
<!-- ... -->
 <dataRef key="teidata.enumerated"/>
<!-- ... -->
</schemaSpec>
```

*Example*

```
<schemaSpec ident="myTEI"
 source="http://www.tei-c.org/Vault/P5/current/xml/tei/odd/p5subset.xml">
<!-- ... -->
 <dataRef name="float"/>
<!-- ... -->
</schemaSpec>
```

*Example*

```
<dataSpec ident="data.integerExample">
  <desc>datatype used for attributes taking an integer value between 0 and
99</desc>
  <content>
    <dataRef name="nonNegativeInteger"
      restriction="[0-9][0-9]?"/>
  </content>
</dataSpec>
```

*Schematron* <sch:rule context="tei:dataRef[tei:dataFacet]"> <sch:assert test="@name" role="nonfatal">Data facets can only be specified for references to datatypes specified by XML Schema Part 2: Datatypes Second Edition — that is, for there to be a 'dataFacet' child there must be a @name attribute.</sch:assert> <sch:report test="@restriction" role="nonfatal">Data facets and restrictions cannot both be expressed on the same data reference — that is, the @restriction attribute cannot be used when a 'dataFacet' element is present.</sch:report> </sch:rule>

*Schematron* <sch:rule context="tei:dataRef[@restriction]"> <sch:assert test="@name" role="nonfatal">Restrictions can only be specified for references to datatypes specified by XML Schema Part 2: Datatypes Second Edition — that is, for there to be a @restriction attribute there must be a @name attribute, too.</sch:assert> </sch:rule>

*Content model*

```
<content>
 <elementRef minOccurs="0"
  maxOccurs="unbounded" key="dataFacet"/>
</content>
```

*Schema Declaration*

```
element dataRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      attribute key { text }?
    | attribute name { text }?
    | attribute ref { text }?
   ),
   attribute restriction { text }?,
   dataFacet*
}
```

---

## <dataSpec>

(datatype specification) documents a datatype. [*23.3. Specification Elements 23.7. Macro Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.identified
  - *@ident*
  - *@predeclare*
  - *@module*

- att.combinable
  - *@mode*

- att.deprecated
  - *@validUntil*

*Member of* model.oddDecl

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*tagdocs:* schemaSpec specGrp

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc gloss

*tagdocs:* altIdent constraintSpec content equiv exemplum listRef remarks valList

*Example*

```
<dataSpec ident="teidata.pointer">
  <desc versionDate="2013-01-19"
    xml:lang="en">defines the range of
     attribute values used to provide a single URI, absolute or relative,
     pointing to some other resource, either within the current document
     or elsewhere.</desc>
  <content>
    <dataRef name="anyURI"/>
  </content>
</dataSpec>
```

*Schematron* <sch:rule role="warn" context="tei:dataSpec/tei:content"> <sch:report test=".//tei:anyElement | .//tei:classRef | .//tei:elementRef"> A datatype specification should not refer to an element or a class. </sch:report> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identSynonyms"/>
   <classRef key="model.descLike"/>
  </alternate>
  <alternate minOccurs="0" maxOccurs="1">
   <elementRef key="content"/>
   <elementRef key="valList"/>
  </alternate>
  <elementRef key="constraintSpec"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="exemplum" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="remarks" minOccurs="0"
```

```
   maxOccurs="unbounded"/>
  <elementRef key="listRef" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element dataSpec
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.identified.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   (
      ( model.identSynonyms | model.descLike )*,
      ( content | valList )?,
      constraintSpec*,
      exemplum*,
      remarks*,
      listRef*
   )
}
```

---

## <datatype>

(datatype) specifies the declared value for an attribute, by referring to any datatype defined by the chosen schema language. [*1.4.2. Datatype Specifications 23.5.3. Attribute List Specification*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@minOccurs` (minimum number of occurrences) indicates the minimum number of times this datatype may occur in an instance of the attribute being defined.
*Derived from* att.repeatable
*Status* Optional
*Datatype* teidata.count
*Default* 1

`@maxOccurs` (maximum number of occurrences) indicates the maximum number of times this datatype may occur in an instance of the attribute being defined.
*Derived from* att.repeatable
*Status* Optional
*Datatype* teidata.unboundedCount
*Default* 1

*Contained by:* attDef
*May contain* ANY

*Note* In the TEI scheme, most datatypes are expressed using pre-defined TEI macros, which map a name in the form `teidata.xxxx` to a datatype defined by RELAX NG or the W3C Schema Language.

*Example*

```
<datatype>
  <dataRef name="token"/>
</datatype>
```

*Example*

```
<datatype>
  <dataRef key="teidata.enumerated"/>
</datatype>
```

*Example* The encoding in the following example requires that the attribute being defined contain at least two URIs in its value, as is the case for the *target* attribute of <join>.

```
<datatype  minOccurs="2"
  maxOccurs="unbounded">
  <dataRef  key="teidata.pointer"/>
</datatype>
```

*Content model*

```
<content>
 <alternate>
  <elementRef key="dataRef"/>
  <anyElement require="http://relaxng.org/ns/structure/1.0"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element datatype
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute minOccurs { text }?,
   attribute maxOccurs { text }?,
   ( dataRef | anyElement_datatype_6 )
}
```

---

## <date>

(date) contains a date in any format. [*3.6.4. Dates and Times 2.2.4. Publication, Distribution, Licensing, etc. 2.6. The Revision Description 3.12.2.4. Imprint, Size of a Document, and Reprint Information 16.2.3. The Setting Description 14.4. Dates*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.calendarSystem
  - *@calendar*
- att.canonical
  - *@key*
  - *@ref*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*

    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
  - att.dimensions
    - *@unit*
    - *@quantity*
    - *@extent*
    - *@precision*
    - *@scope*
    - att.ranging
      * *@atLeast*
      * *@atMost*
      * *@min*
      * *@max*
      * *@confidence*
  - att.duration
    - att.duration.iso
      * *@dur-iso*
    - att.duration.w3c
      * *@dur*
  - att.editLike
    - *@evidence*
    - *@instant*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.dateLike model.publicationStmtPart.detail

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine analytic author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose setting

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal publicationStmt rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* recording u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q ref reg rs ruby sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<date when="1980-02">early February 1980</date>
```

*Example*

```
Given on the <date when="1977-06-12">Twelfth Day
of June in the Year of Our Lord One Thousand Nine Hundred and Seventy-seven of
```

```
     the Republic
     the Two Hundredth and first and of the University the Eighty-Sixth.</date>
```

*Example*

```
     <date  when="1990-09">September 1990</date>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element date
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.calendarSystem.attributes,
   att.canonical.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   att.editLike.attributes,
   att.typed.attributes,
   ( text | model.gLike | model.phrase | model.global )*
}
```

---

## <dateline>

(dateline) contains a brief description of the place, date, time, etc. of production of a letter, newspaper story, or other work, prefixed or suffixed to it as a kind of heading or trailer. [*4.2.2. Openers and Closers*]

*Module* textstructure — Default Text Structure

*Attributes*
- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.written
  - *@hand*

*Member of* model.divWrapper model.pLike.front

*Contained by*

*cmc:* post

*core:* lg list

*drama:* castList epilogue performance prologue

*figures:* figure table

*textstructure:* back body closer div div1 div2 div3 div4 div5 div6 div7 front group opener

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice corr date del distinct ellipsis email emph expan foreign gap gb

gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q ref reg rs ruby sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* docDate

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<dateline>Walden, this 29. of August 1592</dateline>
```

*Example*

```
<div  type="chapter">
  <p>
<!-- ... --> and his heart was going like mad and yes I said yes I will
Yes.</p>
  <closer>
    <dateline>
      <name  type="place">Trieste-Zürich-Paris,</name>
      <date>1914–1921</date>
    </dateline>
  </closer>
</div>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.global"/>
  <elementRef key="docDate"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element dateline
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.written.attributes,
   ( text | model.gLike | model.phrase | model.global | docDate )*
}
```

---

**<death>** (death) contains information about a person's death, such as its date and place. [*16.2.2. The Participant Description*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*
  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.editLike
  - *@evidence*

- *@instant*

- att.naming
  - *@role*
  - *@nymRef*

  - att.canonical
    - *@key*
    - *@ref*

- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **proclaimed**
**assumed**
**verified**
**clinical**
**brain**
**natural**
**unnatural**
**fragmentation**
**dissolution**
*Note* This attribute is not intended to express the cause of death.

*Member of* model.personPart
*Contained by*
*namesdates:* person personGrp persona
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit
*dictionaries:* lang oRef pRef
*figures:* figure formula notatedMusic
*gaiji:* g
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<death  when="1902-10-01"/>
```

*Example*

```
<death  when="1960-12-10">Passed away near
<name  type="place">Aix-la-Chapelle</name>, after suffering from cerebral
palsy. </death>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element death
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.phraseSeq}
```

---

**<decoDesc>** (decoration description) contains a description of the decoration of a manuscript or other object, either as in paragraphs, or as one or more <decoNote> elements. [*11.7.3. Bindings, Seals, and Additional Material*]

*Module* msdescription — Manuscript Description
*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  * *@ana*
- – att.global.change
  * *@change*
- – att.global.facs
  * *@facs*
- – att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- – att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- – att.global.responsibility
  * *@cert*
  * *@resp*
- – att.global.source
  * *@source*

*Member of* model.physDescPart
*Contained by*
*msdescription:* physDesc
*May contain*
*core:* p
*linking:* ab
*msdescription:* decoNote summary
*Example*

```
<decoDesc>
  <p>The start of each book of the Bible with a 10-line historiated
     illuminated initial; prefaces decorated with 6-line blue initials with red
```

```
        penwork flourishing; chapters marked by 3-line plain red initials; verses
        with 1-line initials, alternately blue or red.</p>
    </decoDesc>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="summary" minOccurs="0"/>
   <elementRef key="decoNote" minOccurs="1"
    maxOccurs="unbounded"/>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element decoDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.pLike+ | ( summary?, decoNote+ ) )
}
```

---

## <decoNote>

**<decoNote>** (note on decoration) contains a note describing either a decorative component of a manuscript or other object, or a fairly homogenous class of such components. [*11.7.3. Bindings, Seals, and Additional Material*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs

    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.biblPart model.msItemPart

*Contained by*

*core:* bibl

*msdescription:* binding bindingDesc decoDesc msItem msItemStruct seal sealDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink

objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<decoDesc>
  <decoNote  type="initial">
    <p>The start of each book of the Bible with
        a 10-line historiated illuminated initial;
        prefaces decorated with 6-line blue initials
        with red penwork flourishing; chapters marked by
        3-line plain red initials; verses with 1-line initials,
        alternately blue or red.</p>
  </decoNote>
</decoDesc>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element decoNote
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   macro.specialPara}
```

---

## <def> (definition) contains definition text in a dictionary entry. [*10.3.3.1. Definitions*]

*Module* dictionaries — Dictionaries

*Attributes* • att.global

- *@xml:id*
- *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*
- att.lexicographic.normalized
  - *@norm*
  - *@orig*

*Member of* model.entryPart model.entryPart.top

*Contained by*

*core:* cit

*dictionaries:* dictScrap entry entryFree etym hom re sense

*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<entry>
  <form>
    <orth>competitor</orth>
    <hyph>com|peti|tor</hyph>
    <pron>k@m"petit@(r)</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
  <def>person who competes.</def>
</entry>
```

*Content model*

```
<content>
```

```
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element def
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   macro.paraContent}
```

---

**<default>** (default feature value) represents the value part of a feature-value specification which contains a defaulted value. [*19.9. Default Values*]

*Module* iso-fs — Feature Structures

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.featureVal.single

*Contained by*

*iso-fs:* f fvLib if vAlt vColl vDefault vLabel vMerge vNot vRange

*May contain* Empty element

*Example*

```
<f name="gender">
  <default/>
</f>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element default
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   empty
}
```

---

## <defaultVal>

(default value) specifies the default declared value for an attribute. [*23.5.3. Attribute List Specification*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.combinable
  - *@mode*

- att.deprecated
  - *@validUntil*

*Contained by* tagdocs: attDef
*May contain* Character data only
*Note* any legal declared value or TEI-defined keyword
*Example*

```
<defaultVal>#IMPLIED</defaultVal>
```

*Content model* `<content> <textNode/></content>`
*Schema Declaration*

```
element defaultVal
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
```

```
   text
}
```

---

**<del>** (deletion) contains a letter, word, or passage deleted, marked as deleted, or otherwise indicated as superfluous or spurious in the copy text by an author, scribe, or a previous annotator or corrector. [*3.5.3. Additions, Deletions, and Omissions*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.dimensions
  - *@unit*

- *@quantity*
- *@extent*
- *@precision*
- *@scope*

- att.ranging
  - *@atLeast*
  - *@atMost*
  - *@min*
  - *@max*
  - *@confidence*
- att.transcriptional
  - *@status*
  - *@cause*
  - *@seq*

- att.editLike
  - *@evidence*
  - *@instant*
- att.placement
  - *@place*
- att.written
  - *@hand*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.linePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename

genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw line metamark mod restore retrace secl subst supplied surplus zone

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* This element should be used for deletion of shorter sequences of text, typically single words or phrases. The <delSpan> element should be used for longer sequences of text, for those containing structural subdivisions, and for those containing overlapping additions and deletions.

The text deleted must be at least partially legible in order for the encoder to be able to transcribe it (unless it is restored in a <supplied> tag). Illegible or lost text within a deletion may be marked using the <gap> tag to signal that text is present but has not been transcribed, or is no longer visible. Attributes on the <gap> element may be used to indicate how much text is omitted, the reason for omitting it, etc. If text is not fully legible, the <unclear> element (available when using the additional tagset for transcription of primary sources) should be used to signal the areas of text which cannot be read with confidence in a similar way.

Degrees of uncertainty over what can still be read, or whether a deletion was intended may be indicated by use of the <certainty> element (see *22. Certainty, Precision, and Responsibility*).

There is a clear distinction in the TEI between <del> and <surplus> on the one hand and <gap> or <unclear> on the other. <del> indicates a deletion present in the source being transcribed, which states the author's or a later scribe's intent to cancel or remove text. <surplus> indicates material present in the source being transcribed which should have been so deleted, but which is not in fact. <gap> or <unclear>, by contrast, signal an editor's or encoder's decision to omit something or their inability to read the source text. See sections *12.3.1.7. Text Omitted from or Supplied in the Transcription* and *12.3.3.2. Use of the gap, del, damage, unclear, and supplied Elements in Combination* for the relationship between these and other related elements used in detailed transcription.

*Example*

```
<l>
  <del  rend="overtyped">Mein</del> Frisch
<del  rend="overstrike"  type="primary">schwebt</del>
weht der Wind
</l>
```

*Example*

```
<del  rend="overstrike">
  <gap  reason="illegible"  quantity="5"
    unit="character"/>
</del>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element del
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.transcriptional.attributes,
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   att.typed.attributes,
   macro.paraContent}
```

**<delSpan>** (deleted span of text) marks the beginning of a longer sequence of text deleted, marked as deleted, or otherwise signaled as superfluous or spurious by an author, scribe, annotator, or corrector. [*12.3.1.4. Additions and Deletions*]

*Module* transcr — Representation of Primary Sources

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*
  - att.ranging
    - *@atLeast*

    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
  - att.spanning
    - *@spanTo*
  - att.transcriptional
    - *@status*
    - *@cause*
    - *@seq*
    - att.editLike
      * *@evidence*
      * *@instant*
    - att.placement
      * *@place*
    - att.written
      * *@hand*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.global.edit

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Note* Both the beginning and ending of the deleted sequence must be marked: the beginning by the <delSpan> element, the ending by the target of the *spanTo* attribute.

The text deleted must be at least partially legible, in order for the encoder to be able to transcribe it. If it is not legible at all, the <delSpan> tag should not be used. Rather, the <gap> tag should be employed to signal that text cannot be transcribed, with the value of the *reason* attribute giving the cause for the omission from the transcription as deletion. If it is not fully legible, the <unclear> element should be used to signal the areas of text which cannot be read with confidence. See further sections *12.3.1.7. Text Omitted from or Supplied in the Transcription* and, for the close association of the <delSpan> tag with the <gap>, <damage>, <unclear> and <supplied> elements, *12.3.3.2. Use of the gap, del, damage, unclear, and supplied Elements in Combination.*

The <delSpan> tag should not be used for deletions made by editors or encoders. In these cases, either the <corr> tag or the <gap> tag should be used.

*Example*

```
<p>Paragraph partially deleted. This is the undeleted
portion <delSpan  spanTo="#a23"/>and this the deleted
portion of the paragraph.</p>
<p>Paragraph deleted together with adjacent material.</p>
<p>Second fully deleted paragraph.</p>
<p>Paragraph partially deleted; in the middle of this
paragraph the deletion ends and the anchor point marks
the resumption <anchor  xml:id="a23"/> of the text. ...</p>
```

*Schematron* <sch:rule context="tei:delSpan"> <sch:assert test="@spanTo">The @spanTo attribute of <sch:name/> is required.</sch:assert> </sch:rule>

*Schematron* <sch:rule context="tei:delSpan"> <sch:assert test="@spanTo">L'attribut spanTo est requis.</sch:assert> </sch:rule>

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element delSpan
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.spanning.attributes,
   att.transcriptional.attributes,
```

```
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   att.typed.attributes,
   empty
}
```

---

**<depth>** (depth) contains a measurement from the front to the back of an object, perpendicular to the measurement given by the <width> element. [*11.3.4. Dimensions*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.dimensions
  - *@unit*

- *@quantity*
- *@extent*
- *@precision*
- *@scope*

- att.ranging
  * *@atLeast*
  * *@atMost*
  * *@min*
  * *@max*
  * *@confidence*

*Member of* model.dimLike model.measureLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure measureGrp meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote dimensions explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*gaiji:* g

character data

*Example*

```
<depth unit="in" quantity="4"/>
```

*Content model* `<content> <macroRef key="macro.xtext"/></content>`

*Schema Declaration*

```
element depth
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   macro.xtext}
```

---

## <derivation>

(derivation) describes the nature and extent of originality of this text. [*16.2.1. The Text Description*]

*Module* corpus — Language Corpora

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

  - * *@style*
  - * *@rendition*
  - – att.global.responsibility
  - * *@cert*
  - * *@resp*
  - – att.global.source
  - * *@source*
- att.typed
  - – ~~type~~
  - – @subtype

`@type` categorizes the derivation of the text.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **original** text is original

**revision** text is a revision of some other text

**translation** text is a translation of some other text

**abridgment** text is an abridged version of some other text

**plagiarism** text is plagiarized from some other text

**traditional** text has no obvious source but is one of a number derived from some common ancestor

*Member of* model.textDescPart

*Contained by*

*corpus:* textDesc

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data

*Note* For derivative texts, details of the ancestor may be included in the source description.

*Example*

```
<derivation type="original"/>
```

*Example*

```
<derivation type="translation"
  source="#rosette"/>
<!-- ... -->
<!-- in the sourceDesc: -->
<bibl xml:id="rosette">
  <author>de Béranger, Pierre-Jean</author>. <date>1839</date>.
"<title level="a">Rosette</title>". In <editor>H. Fournier</editor>, ed.
<title level="m">Œuvres complètes de Béranger</title>.
<biblScope unit="volume">Vol 2</biblScope> (p.
<biblScope unit="page">29-30</biblScope>).
</bibl>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element derivation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.phraseSeq.limited}
```

---

**<desc>** (description) contains a short description of the purpose, function, or use of its parent element, or when the parent is a documentation element, describes or defines the object being documented. [*23.4.1. Description of Components*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

  * *@ana*
  - att.global.change
  * *@change*
  - att.global.facs
  * *@facs*
  - att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
  - att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
  - att.global.responsibility
  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.cmc
  - *@generatedBy*
- att.translatable
  - *@versionDate*
- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **deprecationInfo** (deprecation information) This element describes why or how its parent element is being deprecated, typically including recommendations for alternate encoding.

```
<dataSpec module="tei"
  ident="teidata.point"
  validUntil="2050-02-25">
  <desc type="deprecationInfo"
    versionDate="2018-09-14"
    xml:lang="en">Several standards bodies, including NIST in the
USA,
     strongly recommend against ending the representation of a
```

```
number
    with a decimal point. So instead of <q>3.</q> use either
<q>3</q>
    or <q>3.0</q>.</desc>
<!-- ... -->
</dataSpec>
```

*Member of* model.descLike model.labelLike

*Contained by*

*analysis:* interp interpGrp spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* add corr del desc ellipsis emph gap graphic head hi item l lg list listBibl media meeting note noteGrp orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure notatedMusic

*gaiji:* char charDecl glyph

*header:* application category change citeStructure handNote licence listChange listPrefixDef rendition schemaRef scriptNote tagUsage taxonomy unitDef

*iso-fs:* fDescr fsDescr

*linking:* ab altGrp join joinGrp linkGrp listAnnotation seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* climate event listEvent listNym listObject listOrg listPerson listPlace listRelation location occupation org place population relation state terrain trait

*nets:* listForest

*spoken:* incident kinesic transcriptionDesc u vocal writing

*tagdocs:* attDef classSpec constraintDecl constraintSpec dataSpec elementSpec listRef macroSpec model modelGrp modelSequence moduleSpec paramSpec schemaSpec valItem

*textcrit:* lem listApp listWit rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage listTranspose metamark mod restore retrace secl space substJoin supplied surface surplus

*verse:* rhyme

*May contain*

*core:* abbr address bibl biblStruct choice cit date desc distinct email emph expan foreign gloss hi label list listBibl measure measureGrp mentioned name num ptr q quote ref rs said soCalled stage term time title unit

*dictionaries:* lang

*drama:* camera caption castList move sound tech view

*figures:* table

*header:* biblFull idno

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specGrp specGrpRef tag val

*textcrit:* listApp listWit

*textstructure:* floatingText

*transcr:* am ex subst

character data

*Note* When used in a specification element such as <elementSpec>, TEI convention requires that this be expressed as a finite clause, begining with an active verb.

*Example* Example of a <desc> element inside a documentation element.

```
<dataSpec module="tei"
  ident="teidata.point">
  <desc versionDate="2010-10-17"
    xml:lang="en">defines the data type used to express a point in cartesian
space.</desc>
  <content>
    <dataRef name="token"
      restriction="(-?[0-9]+(\.[0-9]+)?,-?[0-9]+(\.[0-9]+)?)"/>
  </content>
<!-- ... -->
</dataSpec>
```

*Example* Example of a <desc> element in a non-documentation element.

```
<place xml:id="KERG2">
  <placeName>Kerguelen Islands</placeName>
<!-- ... -->
  <terrain>
    <desc>antarctic tundra</desc>
  </terrain>
<!-- ... -->
</place>
```

*Schematron* A <desc> with a *type* of deprecationInfo should only occur when its parent element is being deprecated. Furthermore, it should always occur in an element that is being deprecated when <desc> is a valid child of that element. <sch:rule context="tei:desc[ @type eq 'deprecationInfo']"> <sch:assert test="../@validUntil">Information about a deprecation should only be present in a specification element that is being deprecated: that is, only an element that has a @validUntil attribute should have a child <desc type="deprecationInfo">.</sch:assert> </sch:rule>

*Content model*

```
<content>
 <macroRef key="macro.limitedContent"/>
</content>
```

*Schema Declaration*

```
element desc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
```

```
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.translatable.attributes,
   att.typed.attribute.subtype,
   attribute type { "deprecationInfo" }?,
   macro.limitedContent}
```

---

**<dictScrap>** (dictionary scrap) encloses a part of a dictionary entry in which other phrase-level dictionary elements are freely combined. [*10.1. Dictionary Body and Overall Structure 10.2. The Structure of Dictionary Entries*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Member of* model.entryPart.top

*Contained by*

*dictionaries:* entry hom re sense superEntry

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index label lb list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* case colloc def etym form gen gram gramGrp hom hyph iType lang lbl mood number oRef orth pRef per pos pron re sense subc superEntry syll tns usg xr

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May contain any dictionary elements in any combination.

This element is used to mark part of a dictionary entry in which lower level dictionary elements appear, but which does not itself form an identifiable structural unit.

*Example*

```
<entry>
  <dictScrap>
    <orth>biryani</orth> or <orth>biriani</orth>
    <pron>(%bIrI"A:nI)</pron>
    <def>any of a variety of Indian dishes ...</def>
    <etym>[from <lang>Urdu</lang>]</etym>
  </dictScrap>
</entry>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
```

```
  <classRef key="model.gLike"/>
  <classRef key="model.entryPart"/>
  <classRef key="model.morphLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element dictScrap
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      text
    | model.gLike    | model.entryPart    | model.morphLike    | model.phrase    | model.i
}
```

---

**<dim>** contains any single measurement forming part of a dimensional specification of some sort. [*11.3.4. Dimensions*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*
  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.measureLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure measureGrp meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote dimensions explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*gaiji:* g

character data

*Note* The specific elements <width>, <height>, and <depth> should be used in preference to this generic element wherever appropriate.

*Example*

```
<dim type="circumference" extent="4.67 in"/>
```

*Content model* `<content> <macroRef key="macro.xtext"/></content>`

*Schema Declaration*

```
element dim
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.typed.attributes,
   macro.xtext}
```

---

## <dimensions> (dimensions) contains a dimensional specification. [*11.3.4. Dimensions*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.typed
  - ~~type~~
  - @subtype

`@type` indicates which aspect of the object is being measured.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **leaves** dimensions relate to one or more leaves (e.g. a single leaf, a gathering, or a separately bound part)

**ruled** dimensions relate to the area of a leaf which has been ruled in preparation for writing.

**pricked** dimensions relate to the area of a leaf which has been pricked out in preparation for ruling (used where this differs significantly from the ruled area, or where the ruling is not measurable).

**written** dimensions relate to the area of a leaf which has been written, with the height measured from the top of the minims on the top line of writing, to the bottom of the minims on the bottom line of writing.

**miniatures** dimensions relate to the miniatures within the manuscript

**binding** dimensions relate to the binding in which the codex or manuscript is contained

**box** dimensions relate to the box or other container in which the manuscript is stored.

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*msdescription:* depth dim height width

*Note* Contains no more than one of each of the specialized elements used to express a three-dimensional object's height, width, and depth, combined with any number of other kinds of dimensional specification.

*Example*

```
<dimensions type="leaves">
  <height scope="range">157-160</height>
  <width>105</width>
</dimensions>
<dimensions type="ruled">
  <height scope="most">90</height>
  <width scope="most">48</width>
</dimensions>
<dimensions unit="in">
  <height>12</height>
  <width>10</width>
</dimensions>
```

*Example* This element may be used to record the dimensions of any text-bearing object, not necessarily a codex. For example:

```
<dimensions type="panels">
  <height scope="all">7004</height>
  <width scope="all">1803</width>
  <dim type="relief" unit="mm">345</dim>
</dimensions>
```

This might be used to show that the inscribed panels on some (imaginary) monument are all the same size (7004 by 1803 cm) and stand out from the rest of the monument by 345 mm.

*Example* When simple numeric quantities are involved, they may be expressed on the *quantity* attribute of any or all of the child elements, as in the following example:

```
<dimensions type="leaves">
  <height scope="range">157-160</height>
  <width quantity="105"/>
</dimensions>
<dimensions type="ruled">
  <height unit="cm" scope="most"
    quantity="90"/>
  <width unit="cm" scope="most" quantity="48"/>
</dimensions>
<dimensions unit="in">
  <height quantity="12"/>
  <width quantity="10"/>
</dimensions>
```

*Schematron* <sch:rule context="tei:dimensions"> <sch:report test="count(tei:width) gt 1"> The element <sch:name/> may appear once only </sch:report> <sch:report test="count(tei:height) gt 1"> The element <sch:name/> may appear once only </sch:report> <sch:report test="count(tei:depth) gt 1"> The element <sch:name/> may appear once only </sch:report> </sch:rule>

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
```

```
  <elementRef key="dim"/>
  <classRef key="model.dimLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element dimensions
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   ( dim | model.dimLike )*
}
```

---

**<distinct>** identifies any word or phrase which is regarded as linguistically distinct, for example as archaic, technical, dialectal, non-preferred, etc., or as forming part of a sublanguage. [*3.3.2.3. Other Linguistically Distinct Material*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.typed
  - ~~type~~
  - @subtype

`@type` specifies the sublanguage or register to which the word or phrase is being assigned.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated

`@time` specifies how the phrase is distinct diachronically.
*Status* Optional
*Datatype* teidata.text

`@space` specifies how the phrase is distinct diatopically.
*Status* Optional
*Datatype* teidata.text

`@social` specifies how the phrase is distinct diastratically.
*Status* Optional
*Datatype* teidata.text

*Member of* model.emphLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
Next morning a boy
in that dormitory confided to his bosom friend, a
<distinct type="ps_slang">fag</distinct> of
Macrea's, that there was trouble in their midst which King
<distinct type="archaic">would fain</distinct>
```

```
keep secret.
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element distinct
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   attribute time { text }?,
   attribute space { text }?,
   attribute social { text }?,
   macro.phraseSeq}
```

---

## <distributor>

(distributor) supplies the name of a person or other agency responsible for the distribution of a text. [*2.2.4. Publication, Distribution, Licensing, etc.*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*

    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.canonical
  - *@key*
  - *@ref*

*Member of* model.imprintPart model.publicationStmtPart.agency

*Contained by*

*core:* bibl imprint

*header:* publicationStmt

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme
character data

*Example*

```
<distributor>Oxford Text Archive</distributor>
<distributor>Redwood and Burn Ltd</distributor>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element distributor
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   macro.phraseSeq}
```

---

## <district>

(district) contains the name of any kind of subdivision of a settlement, such as a parish, ward, or other administrative or geographic unit. [*14.2.3. Place Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.naming
  - *@role*
  - *@nymRef*

  - att.canonical
    - *@key*
    - *@ref*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.placeNamePart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage unitDef

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions altIdentifier catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msIdentifier musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectIdentifier objectName occupation offset org orgName persName persPronouns place placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<placeName>
  <district  type="ward">Jericho</district>
  <settlement>Oxford</settlement>
</placeName>
```

*Example*

```
<placeName>
  <district  type="area">South Side</district>
  <settlement>Chicago</settlement>
</placeName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element district
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.naming.attributes,
   att.canonical.attributes,
```

```
   att.typed.attributes,
   macro.phraseSeq}
```

---

** ** (text division) contains a subdivision of the front, body, or back of a text. [*4.1. Divisions of the Body*]

*Module* textstructure — Default Text Structure

*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

• att.declaring
- *@decls*

• att.divLike
- *@org*
- *@sample*

  - att.fragmentable
    * *@part*
  - att.metrical
    * *@met*
    * *@real*
    * *@rhyme*
- att.placement
  - *@place*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*

*Member of* model.divLike

*Contained by*

*textcrit:* lem rdg

*textstructure:* back body div front

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct cb cit desc divGen ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition schemaSpec specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline div docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Example*

```
<body>
  <div  type="part">
    <head>Fallacies of Authority</head>
    <p>The subject of which is Authority in various shapes, and the object, to
repress all
        exercise of the reasoning faculty.</p>
```

```
   <div  n="1"  type="chapter">
     <head>The Nature of Authority</head>
     <p>With reference to any proposed measures having for their object the
greatest
          happiness of the greatest number [...]</p>
     <div  n="1.1"  type="section">
       <head>Analysis of Authority</head>
       <p>What on any given occasion is the legitimate weight or influence to
be attached to
             authority [...] </p>
     </div>
     <div  n="1.2"  type="section">
       <head>Appeal to Authority, in What Cases Fallacious.</head>
       <p>Reference to authority is open to the charge of fallacy when [...]
</p>
     </div>
   </div>
  </div>
</body>
```

*Schematron* <sch:rule context="tei:l//tei:div"> <sch:assert test="ancestor::tei:floatingText"> Abstract model violation: Metrical lines may not contain higher-level structural elements such as div, unless div is a descendant of floatingText. </sch:assert> </sch:rule>

*Schematron* <sch:rule context="tei:div"> <sch:report test="(ancestor::tei:p or ancestor::tei:ab) and not(ancestor::tei:floatingText)"> Abstract model violation: p and ab may not contain higher-level structural elements such as div, unless div is a descendant of floatingText. </sch:report> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="0">
   <alternate>
    <sequence minOccurs="1"
     maxOccurs="unbounded">
     <alternate>
      <classRef key="model.divLike"/>
      <classRef key="model.divGenLike"/>
     </alternate>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
    <sequence>
     <sequence minOccurs="1"
      maxOccurs="unbounded">
      <alternate minOccurs="1"
       maxOccurs="1">
       <elementRef key="schemaSpec"/>
       <classRef key="model.common"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
```

```
     <sequence minOccurs="0"
      maxOccurs="unbounded">
      <alternate>
       <classRef key="model.divLike"/>
       <classRef key="model.divGenLike"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
    </sequence>
   </alternate>
   <sequence minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.divBottom"/>
    <classRef key="model.global"
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element div
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.divLike.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.placement.attributes,
   att.typed.attributes,
   att.written.attributes,
   (
      ( model.divTop | model.global )*,
      (
         (
            (
               ( ( ( model.divLike | model.divGenLike ), model.global* )+ )
             | (
                  ( ( ( schemaSpec | model.common ), model.global* )+ ),
                  ( ( ( model.divLike | model.divGenLike ), model.global* )* )
               )
            ),
            ( ( model.divBottom, model.global* )* )
         )?
      )
   )
}
```

**<div1>** (level-1 text division) contains a first-level subdivision of the front, body, or back of a text. [*4.1.2. Numbered Divisions*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declaring
  - *@decls*
- att.divLike
  - *@org*
  - *@sample*
  - att.fragmentable
    - *@part*

- att.metrical
  - *@met*
  - *@real*
  - *@rhyme*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.div1Like

*Contained by*

*textstructure:* back body front

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct cb cit desc divGen ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition schemaSpec specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline div2 docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* any sequence of low-level structural elements, possibly grouped into lower subdivisions.

*Example*

```
<div1 xml:id="levi" n="I" type="part">
  <head>Part I: Of Man </head>
  <div2 xml:id="levi1" n="1" type="chapter">
    <head>Chap. I. Of Sense </head>
    <p>Concerning the Thoughts of man... </p>
  </div2>
</div1>
<div1 xml:id="levii" n="II" type="part">
  <head>Part II: Of Common-Wealth</head>
</div1>
```

*Content model*

```
<content>
 <sequence>
```

```
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="0">
   <alternate>
    <sequence minOccurs="1"
     maxOccurs="unbounded">
     <alternate>
      <classRef key="model.div2Like"/>
      <classRef key="model.divGenLike"/>
     </alternate>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
    <sequence>
     <sequence minOccurs="1"
      maxOccurs="unbounded">
      <alternate minOccurs="1"
       maxOccurs="1">
       <elementRef key="schemaSpec"/>
       <classRef key="model.common"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
     <sequence minOccurs="0"
      maxOccurs="unbounded">
      <alternate>
       <classRef key="model.div2Like"/>
       <classRef key="model.divGenLike"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
    </sequence>
   </alternate>
   <sequence minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.divBottom"/>
    <classRef key="model.global"
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element div1
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
```

```
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.divLike.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   (
      ( model.divTop | model.global )*,
      (
         (
            (
               ( ( ( model.div2Like | model.divGenLike ), model.global* )+ )
             | (
                  ( ( ( schemaSpec | model.common ), model.global* )+ ),
                  ( ( ( model.div2Like | model.divGenLike ), model.global* )* )
               )
            ),
            ( ( model.divBottom, model.global* )* )
         )?
      )
   )
}
```

---

**<div2>** (level-2 text division) contains a second-level subdivision of the front, body, or back of a text. [*4.1.2. Numbered Divisions*]

*Module* textstructure — Default Text Structure

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.declaring
  - *@decls*
- att.divLike
  - *@org*
  - *@sample*

  - att.fragmentable
    * *@part*
  - att.metrical
    * *@met*
    * *@real*
    * *@rhyme*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.div2Like
*Contained by*
*textstructure:* div1
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*cmc:* post
*core:* bibl biblStruct cb cit desc divGen ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage
*dictionaries:* entry entryFree superEntry
*drama:* camera caption castList move sound spGrp tech view
*figures:* figure notatedMusic table
*header:* biblFull
*iso-fs:* fLib fs fvLib
*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* msDesc
*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation
*nets:* eTree forest graph listForest tree
*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition schemaSpec specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline div3 docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* any sequence of low-level structural elements, possibly grouped into lower subdivisions.

*Example*

```
<div1 n="2" type="part">
  <head>The Second Partition:
     The Cure of Melancholy</head>
  <div2 n="2.1" type="section">
    <div3 n="2.1.1" type="member">
      <div4 n="2.1.1.1" type="subsection">
        <head>Unlawful Cures rejected.</head>
        <p>Inveterate melancholy, howsoever it may seem to
                be a continuate, inexorable disease, hard to be
                cured, accompanying them to their graves most part
                (as <ref target="#a">Montanus</ref> observes), yet many
                times it may be helped...
        </p>
      </div4>
    </div3>
  </div2>
  <div2 n="2.2" type="section">
    <div3 n="2.2.1" type="member">
      <head>Sect. II. Memb. I</head>
      <p/>
    </div3>
  </div2>
  <div2 n="2.3" type="section">
    <div3 n="2.3.1" type="member">
      <head>Sect. III. Memb. I</head>
      <p/>
    </div3>
  </div2>
</div1>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="0">
   <alternate>
    <sequence minOccurs="1"
     maxOccurs="unbounded">
     <alternate>
      <classRef key="model.div3Like"/>
      <classRef key="model.divGenLike"/>
     </alternate>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
```

```
    </sequence>
    <sequence>
     <sequence minOccurs="1"
      maxOccurs="unbounded">
      <alternate minOccurs="1"
       maxOccurs="1">
       <elementRef key="schemaSpec"/>
       <classRef key="model.common"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
     <sequence minOccurs="0"
      maxOccurs="unbounded">
      <alternate>
       <classRef key="model.div3Like"/>
       <classRef key="model.divGenLike"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
    </sequence>
   </alternate>
   <sequence minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.divBottom"/>
    <classRef key="model.global"
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element div2
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.divLike.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   (
      ( model.divTop | model.global )*,
      (
         (
            (
               ( ( ( model.div3Like | model.divGenLike ), model.global* )+ )
             | (
                  ( ( ( schemaSpec | model.common ), model.global* )+ ),
```

```
                  ( ( ( model.div3Like | model.divGenLike ), model.global* )* )
               )
            ),
            ( ( model.divBottom, model.global* )* )
         )?
      )
   )
}
```

---

**<div3>** (level-3 text division) contains a third-level subdivision of the front, body, or back of a text. [*4.1.2. Numbered Divisions*]

*Module* textstructure — Default Text Structure

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declaring

  - *@decls*
- att.divLike
  - *@org*
  - *@sample*
  - att.fragmentable
    - *@part*
  - att.metrical
    - *@met*
    - *@real*
    - *@rhyme*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.div3Like

*Contained by*

*textstructure:* div2

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct cb cit desc divGen ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition schemaSpec specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline div4 docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* any sequence of low-level structural elements, possibly grouped into lower subdivisions.

*Example*

```
<div2 n="2.2" type="section">
  <div3 n="2.2.1" type="member">
    <head>Sect. II. Memb. I</head>
    <p/>
  </div3>
```

```
      <div3  n="2.2.2"  type="member">
        <head>Memb. II Retention and Evacuation rectified.</head>
        <p/>
      </div3>
      <div3  n="2.2.3"  type="member">
        <head>Memb. III Ayr rectified. With a digression of the Ayr.</head>
        <p/>
      </div3>
    </div2>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="0">
   <alternate>
    <sequence minOccurs="1"
     maxOccurs="unbounded">
     <alternate>
      <classRef key="model.div4Like"/>
      <classRef key="model.divGenLike"/>
     </alternate>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
    <sequence>
     <sequence minOccurs="1"
      maxOccurs="unbounded">
      <alternate minOccurs="1"
       maxOccurs="1">
       <elementRef key="schemaSpec"/>
       <classRef key="model.common"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
     <sequence minOccurs="0"
      maxOccurs="unbounded">
      <alternate>
       <classRef key="model.div4Like"/>
       <classRef key="model.divGenLike"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
    </sequence>
   </alternate>
   <sequence minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.divBottom"/>
    <classRef key="model.global"
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
```

```
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element div3
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.divLike.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   (
      ( model.divTop | model.global )*,
      (
         (
            (
               ( ( ( model.div4Like | model.divGenLike ), model.global* )+ )
             | (
                  ( ( ( schemaSpec | model.common ), model.global* )+ ),
                  ( ( ( model.div4Like | model.divGenLike ), model.global* )* )
               )
            ),
            ( ( model.divBottom, model.global* )* )
         )?
      )
   )
}
```

---

**<div4>** (level-4 text division) contains a fourth-level subdivision of the front, body, or back of a text. [*4.1.2. Numbered Divisions*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change

    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.declaring
  - *@decls*
- att.divLike
  - *@org*
  - *@sample*

  - att.fragmentable
    * *@part*
  - att.metrical
    * *@met*
    * *@real*
    * *@rhyme*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.div4Like
*Contained by*
*textstructure:* div3
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*cmc:* post

*core:* bibl biblStruct cb cit desc divGen ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition schemaSpec specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline div5 docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* any sequence of low-level structural elements, possibly grouped into lower subdivisions.

*Example*

```
<div3  n="2.2.1"  type="member">
  <head>Sect. II. Memb. I</head>
  <div4  n="2.2.1.1"  type="subsection">
    <head>Subsect I. — Dyet rectified in substance.</head>
    <p>Diet, <term  xml:lang="grc">diaitotiku</term>,
<term  xml:lang="la">victus</term> or
        living </p>
  </div4>
  <div4  n="2.2.2.1"  type="subsection">
    <head>Subsect II. — Dyet rectified in quantity.</head>
    <p>Man alone, saith Cardan, eates and drinks without appetite, and useth
all his pleasures
       without necessity </p>
  </div4>
</div3>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="0">
   <alternate>
    <sequence minOccurs="1"
     maxOccurs="unbounded">
     <alternate>
      <classRef key="model.div5Like"/>
      <classRef key="model.divGenLike"/>
     </alternate>
```

```
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
    <sequence>
     <sequence minOccurs="1"
      maxOccurs="unbounded">
      <alternate minOccurs="1"
       maxOccurs="1">
       <elementRef key="schemaSpec"/>
       <classRef key="model.common"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
     <sequence minOccurs="0"
      maxOccurs="unbounded">
      <alternate>
       <classRef key="model.div5Like"/>
       <classRef key="model.divGenLike"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
    </sequence>
   </alternate>
   <sequence minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.divBottom"/>
    <classRef key="model.global"
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element div4
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.divLike.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   (
      ( model.divTop | model.global )*,
      (
         (
            (
               ( ( ( model.div5Like | model.divGenLike ), model.global* )+ )
```

```
              | (
                    ( ( ( schemaSpec | model.common ), model.global* )+ ),
                    ( ( ( model.div5Like | model.divGenLike ), model.global* )* )
                )
             ),
             ( ( model.divBottom, model.global* )* )
          )?
       )
    )
}
```

---

**<div5>** (level-5 text division) contains a fifth-level subdivision of the front, body, or back of a text. [*4.1.2. Numbered Divisions*]

*Module* textstructure — Default Text Structure

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source

    * *@source*
  - att.declaring
    – *@decls*
  - att.divLike
    – *@org*
    – *@sample*
    – att.fragmentable
    * *@part*
    – att.metrical
    * *@met*
    * *@real*
    * *@rhyme*
  - att.typed
    – *@type*
    – *@subtype*

*Member of* model.div5Like

*Contained by*

*textstructure:* div4

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct cb cit desc divGen ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition schemaSpec specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline div6 docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* any sequence of low-level structural elements, possibly grouped into lower subdivisions.

*Example*

```
<div2 type="chapter">
  <head>Recipes</head>
```

```
  <head>Chapter VI.</head>
  <div3>
    <head>Fruit and vegetable soups</head>
    <p>...</p>
    <div4>
      <head>Stocks for all kinds of soups</head>
      <div5  type="recipe">
        <head>Rich strong stock</head>
<!-- ... -->
      </div5>
      <div5  type="recipe">
        <head>Medium Stock</head>
<!-- ... -->
      </div5>
    </div4>
    <div4  type="recipe">
      <head>Apple soup</head>
      <div5>
        <head>Ingredients</head>
        <list>
          <item>2 lbs. of good boiling apples,</item>
          <item>3/4 teaspoonful of white pepper,</item>
          <item>6 cloves,</item>
          <item>cayenne or ginger to taste,</item>
          <item>3 quarts of medium stock</item>
        </list>
      </div5>
      <div5>
        <head>Mode</head>
        <p>Peel and quarter the apples taking out their cores; put them into
               the stock, stew them gently till tender, Rub the whole through a
               strainer, add the seasoning. give it one boil up, and serve.</p>
      </div5>
      <div5>
        <head>Time</head>
        <p>1 hour.</p>
      </div5>
      <div5>
        <head>Average cost</head>
        <p>per quart, 1s.</p>
      </div5>
      <div5>
        <head>Seasonable</head>
        <p>from September to December.</p>
      </div5>
      <div5>
        <head>Sufficient</head>
        <p> for 10 persons</p>
      </div5>
      <div5>
        <head>The apple</head>
        <p>This useful fruit is mentioned in Holy Writ; and Homer describes it
               as valuable in his time... As a food, the apple cannot be
considered
               to rank high, as more than the half of it consists of water, and
               the rest of its properties are not the most nourishing. It is
               however a useful adjunct to other kinds of food, and, when
```

```
cooked, is
               esteemed as slightly laxative.</p>
     </div5>
    </div4>
    <div4  type="recipe">
     <head>Artichoke (Jerusalem) soup</head>
     <p>...</p>
    </div4>
<!-- other recipes here -->
  </div3>
</div2>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="0">
   <alternate>
    <sequence minOccurs="1"
     maxOccurs="unbounded">
     <alternate>
      <classRef key="model.div6Like"/>
      <classRef key="model.divGenLike"/>
     </alternate>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
    <sequence>
     <sequence minOccurs="1"
      maxOccurs="unbounded">
      <alternate minOccurs="1"
       maxOccurs="1">
       <elementRef key="schemaSpec"/>
       <classRef key="model.common"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
     <sequence minOccurs="0"
      maxOccurs="unbounded">
      <alternate>
       <classRef key="model.div6Like"/>
       <classRef key="model.divGenLike"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
    </sequence>
   </alternate>
   <sequence minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.divBottom"/>
    <classRef key="model.global"
```

```
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element div5
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.divLike.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   (
      ( model.divTop | model.global )*,
      (
         (
            (
               ( ( ( model.div6Like | model.divGenLike ), model.global* )+ )
             | (
                  ( ( ( schemaSpec | model.common ), model.global* )+ ),
                  ( ( ( model.div6Like | model.divGenLike ), model.global* )* )
               )
            ),
            ( ( model.divBottom, model.global* )* )
         )?
      )
   )
}
```

---

**<div6>** (level-6 text division) contains a sixth-level subdivision of the front, body, or back of a text. [*4.1.2. Numbered Divisions*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

    * *@ana*
  – att.global.change
    * *@change*
  – att.global.facs
    * *@facs*
  – att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  – att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  – att.global.responsibility
    * *@cert*
    * *@resp*
  – att.global.source
    * *@source*
- att.declaring
  – *@decls*
- att.divLike
  – *@org*
  – *@sample*

  – att.fragmentable
    * *@part*
  – att.metrical
    * *@met*
    * *@real*
    * *@rhyme*
- att.typed
  – *@type*
  – *@subtype*

*Member of* model.div6Like
*Contained by*
*textstructure:* div5
*May contain*
*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct cb cit desc divGen ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition schemaSpec specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline div7 docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* any sequence of low-level structural elements, possibly grouped into lower subdivisions.

*Example*

```
<div2  type="chapter">
  <head>Recipes</head>
  <head>Chapter VI.</head>
  <div3>
    <head>Fruit and vegetable soups</head>
    <p>...</p>
    <div4>
      <head>Stocks for all kinds of soups</head>
      <div5  type="recipe">
        <head>Rich strong stock</head>
        <div6>
          <head>Ingredients</head>
          <list>
            <item>4 lbs of shin of beef,</item>
            <item>4 lbs of knuckle of veal,</item>
<!-- ... -->
            <item>4 quarts of water</item>
          </list>
        </div6>
        <div6>
          <head>Mode</head>
          <p>Line a delicately clean stewpan... Strain through a very
                   fine hair sieve, or tammy, and it will be fit for use</p>
        </div6>
        <div6>
          <head>Time</head>
          <p>5 hours.</p>
        </div6>
        <div6>
          <head>Average cost</head>
```

```
          <p>1s 3d. per quart</p>
         </div6>
       </div5>
       <div5  type="recipe">
         <head>Medium Stock</head>
<!-- ... -->
       </div5>
     </div4>
<!-- ... -->
  </div3>
</div2>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="0">
   <alternate>
    <sequence minOccurs="1"
     maxOccurs="unbounded">
     <alternate>
      <classRef key="model.div7Like"/>
      <classRef key="model.divGenLike"/>
     </alternate>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
    <sequence>
     <sequence minOccurs="1"
      maxOccurs="unbounded">
      <alternate minOccurs="1"
       maxOccurs="1">
       <elementRef key="schemaSpec"/>
       <classRef key="model.common"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
     <sequence minOccurs="0"
      maxOccurs="unbounded">
      <alternate>
       <classRef key="model.div7Like"/>
       <classRef key="model.divGenLike"/>
      </alternate>
      <classRef key="model.global"
       minOccurs="0" maxOccurs="unbounded"/>
     </sequence>
    </sequence>
   </alternate>
   <sequence minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.divBottom"/>
    <classRef key="model.global"
```

```
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element div6
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.divLike.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   (
      ( model.divTop | model.global )*,
      (
         (
            (
               ( ( ( model.div7Like | model.divGenLike ), model.global* )+ )
             | (
                  ( ( ( schemaSpec | model.common ), model.global* )+ ),
                  ( ( ( model.div7Like | model.divGenLike ), model.global* )* )
               )
            ),
            ( ( model.divBottom, model.global* )* )
         )?
      )
   )
}
```

---

**<div7>** (level-7 text division) contains the smallest possible subdivision of the front, body or back of a text, larger than a paragraph. [*4.1.2. Numbered Divisions*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

    * *@ana*
  – att.global.change
    * *@change*
  – att.global.facs
    * *@facs*
  – att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  – att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  – att.global.responsibility
    * *@cert*
    * *@resp*
  – att.global.source
    * *@source*
- att.declaring
  – *@decls*
- att.divLike
  – *@org*
  – *@sample*
  – att.fragmentable
    * *@part*
  – att.metrical
    * *@met*
    * *@real*
    * *@rhyme*
- att.typed
  – *@type*
  – *@subtype*

*Member of* model.div7Like
*Contained by*
*textstructure:* div6
*May contain*
*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct cb cit desc ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition schemaSpec specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* any sequence of low-level structural elements, e.g., paragraphs (<p>), lists (<list>), or examples (<eg> or <egXML>).

*Example*

```
<div2 type="chapter">
  <head>Recipes</head>
  <head>Chapter VI.</head>
  <div3>
    <head>Fruit and vegetable soups</head>
    <p>...</p>
    <div4>
      <head>Stocks for all kinds of soups</head>
      <div5 type="recipe">
        <head>Asparagus soup</head>
        <div6 type="altRecipe">
          <head>I.</head>
          <div7>
            <head>Ingredients</head>
            <list>
              <item> ...</item>
            </list>
          </div7>
          <div7>
            <head>Mode</head>
            <p>Put the beef, cut into pieces and rolled in flour, into a
                        stewpan...</p>
          </div7>
<!-- ... -->
        </div6>
        <div6 type="altRecipe">
          <head>II.</head>
          <div7>
            <head>Ingredients</head>
```

```
            <list>
              <item> ...</item>
            </list>
          </div7>
          <div7>
            <head>Mode</head>
            <p>Boil the peas, and rub them through a sieve; add the gravy...</p>
          </div7>
        </div6>
      </div5>
    </div4>
  </div3>
</div2>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="0">
   <sequence minOccurs="1"
    maxOccurs="unbounded">
    <alternate minOccurs="1" maxOccurs="1">
     <elementRef key="schemaSpec"/>
     <classRef key="model.common"/>
    </alternate>
    <classRef key="model.global"
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
   <sequence minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.divBottom"/>
    <classRef key="model.global"
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element div7
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.divLike.attributes,
```

```
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   (
      ( model.divTop | model.global )*,
      (
         (
            ( ( ( schemaSpec | model.common ), model.global* )+ ),
            ( ( model.divBottom, model.global* )* )
         )?
      )
   )
}
```

---

**<divGen>** (automatically generated text division) indicates the location at which a textual division generated automatically by a text-processing application is to appear. [*3.9.2. Index Entries*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*

  * *@resp*
- att.global.source
  * *@source*
- att.typed
- ~~type~~
- @subtype

`@type` specifies what type of generated text division (e.g. index, table of contents, etc.) is to appear.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **index** an index is to be generated and inserted at this point.

**toc** a table of contents

**figlist** a list of figures

**tablist** a list of tables

*Note* Valid values are application-dependent; those shown are of obvious utility in document production, but are by no means exhaustive.

*Member of* model.divGenLike model.frontPart

*Contained by*

*textstructure:* back body div div1 div2 div3 div4 div5 div6 front

*May contain*

*core:* head

*Note* This element is intended primarily for use in document production or manipulation, rather than in the transcription of pre-existing materials; it makes it easier to specify the location of indices, tables of contents, etc., to be generated by text preparation or word processing software.

*Example* One use for this element is to allow document preparation software to generate an index and insert it in the appropriate place in the output. The example below assumes that the *indexName* attribute on <index> elements in the text has been used to specify index entries for the two generated indexes, named NAMES and THINGS:

```
<back>
  <div1  type="backmat">
    <head>Bibliography</head>
<!-- ... -->
  </div1>
  <div1  type="backmat">
    <head>Indices</head>
    <divGen  n="Index Nominum"  type="NAMES"/>
    <divGen  n="Index Rerum"  type="THINGS"/>
  </div1>
</back>
```

*Example* Another use for <divGen> is to specify the location of an automatically produced table of contents:

```
<front>
<!--<titlePage>...</titlePage>-->
  <divGen  type="toc"/>
  <div>
    <head>Preface</head>
    <p> ... </p>
  </div>
</front>
```

*Content model*

```
<content>
 <classRef key="model.headLike"
  minOccurs="0" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element divGen
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   model.headLike*
}
```

---

## <docAuthor>

<docAuthor> (document author) contains the name of the author of the document, as given on the title page (often but not always contained in a byline). [*4.6. Title Pages*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*

    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.cmc
  - *@generatedBy*

*Member of* model.divWrapper model.pLike.front model.titlepagePart

*Contained by*

*cmc:* post

*core:* lg list

*drama:* castList epilogue performance prologue

*figures:* figure table

*msdescription:* msItem

*textcrit:* lem rdg

*textstructure:* back body byline div div1 div2 div3 div4 div5 div6 div7 front group titlePage

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The document author's name often occurs within a byline, but the <docAuthor> element may be used whether the <byline> element is used or not. It should be used only for the author(s) of the entire document, not for author(s) of any subset or part of it. (Attributions of authorship of a subset or part of the document, for example of a chapter in a textbook or an article in a newspaper, may be encoded with <byline> without <docAuthor>.)

*Example*

```
<titlePage>
  <docTitle>
    <titlePart>Travels into Several Remote Nations of the World, in Four
        Parts.</titlePart>
  </docTitle>
  <byline> By <docAuthor>Lemuel Gulliver</docAuthor>, First a Surgeon,
     and then a Captain of several Ships</byline>
</titlePage>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element docAuthor
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.cmc.attributes,
   macro.phraseSeq}
```

---

## <docDate>

(document date) contains the date of a document, as given on a title page or in a dateline. [*4.6. Title Pages*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.calendarSystem
  - *@calendar*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*

  * *@notAfter-iso*
  * *@from-iso*
  * *@to-iso*

– att.datable.w3c
  * *@when*
  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*

*Member of* model.divWrapper model.pLike.front model.titlepagePart

*Contained by*

*cmc:* post

*core:* lg list

*drama:* castList epilogue performance prologue

*figures:* figure table

*msdescription:* msItem

*textcrit:* lem rdg

*textstructure:* back body dateline div div1 div2 div3 div4 div5 div6 div7 docImprint front group titlePage

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Cf. the general <date> element in the core tag set. This specialized element is provided for convenience in marking and processing the date of the documents, since it is likely to require specialized handling for

many applications. It should be used only for the date of the entire document, not for any subset or part of it.

*Example*

```
<docImprint>Oxford, Clarendon Press, <docDate>1987</docDate>
</docImprint>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element docDate
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.calendarSystem.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   macro.phraseSeq}
```

---

## <docEdition> (document edition) contains an edition statement as presented on a title page of a document.

[*4.6. Title Pages*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.pLike.front model.titlepagePart

*Contained by*

*msdescription:* msItem

*textcrit:* lem rdg

*textstructure:* back front titlePage

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Cf. the <edition> element of bibliographic citation. As usual, the shorter name has been given to the more frequent element.

*Example*

```
<docEdition>The Third edition Corrected</docEdition>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element docEdition
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.paraContent}
```

---

## <docImprint>

(document imprint) contains the imprint statement (place and date of publication, publisher name), as given (usually) at the foot of a title page. [*4.6. Title Pages*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.pLike.front model.titlepagePart

*Contained by*

*msdescription:* msItem

*textcrit:* lem rdg

*textstructure:* back front titlePage

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr pubPlace publisher q ref reg rs ruby sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* docDate

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Cf. the <imprint> element of bibliographic citations. As with title, author, and editions, the shorter name is reserved for the element likely to be used more often.

*Example*

```
<docImprint>Oxford, Clarendon Press, 1987</docImprint>
```

Imprints may be somewhat more complex:

```
<docImprint>
  <pubPlace>London</pubPlace>
Printed for <name>E. Nutt</name>,
at
<pubPlace>Royal Exchange</pubPlace>;
<name>J. Roberts</name> in
<pubPlace>wick-Lane</pubPlace>;
<name>A. Dodd</name> without
<pubPlace>Temple-Bar</pubPlace>;
and <name>J. Graves</name> in
<pubPlace>St. James's-street.</pubPlace>
  <date>1722.</date>
</docImprint>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <elementRef key="pubPlace"/>
  <elementRef key="docDate"/>
  <elementRef key="publisher"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element docImprint
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      text
    | model.gLike    | model.phrase    | pubPlace    | docDate    | publisher    | model.g
}
```

---

**<docTitle>** (document title) contains the title of a document, including all its constituents, as given on a title page. [*4.6. Title Pages*]

*Module* textstructure — Default Text Structure

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.canonical
  - *@key*
  - *@ref*

*Member of* model.pLike.front model.titlepagePart

*Contained by*

*msdescription:* msItem

*textcrit:* lem rdg

*textstructure:* back front titlePage

*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* cb ellipsis gap gb index lb milestone note noteGrp pb
*figures:* figure notatedMusic
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*spoken:* incident kinesic pause shift vocal writing
*textcrit:* app witDetail
*textstructure:* titlePart
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
*Example*

```
<docTitle>
  <titlePart  type="main">The DUNCIAD, VARIOURVM.</titlePart>
  <titlePart  type="sub">WITH THE PROLEGOMENA of SCRIBLERUS.</titlePart>
</docTitle>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.global"
   minOccurs="0" maxOccurs="unbounded"/>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="titlePart"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element docTitle
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   ( model.global*, ( ( titlePart, model.global* )+ ) )
}
```

---

**<domain>** (domain of use) describes the most important social context in which the text was realized or for which it is intended, for example private vs. public, education, religion, etc. [*16.2.1. The Text Description*]
*Module* corpus — Language Corpora

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - ~~type~~
  - @subtype

`@type` categorizes the domain of use.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **art** art and entertainment
**domestic** domestic and private
**religious** religious and ceremonial
**business** business and work place
**education** (education)
**govt** (government) government and law

**public** other forms of public context

*Member of* model.textDescPart

*Contained by*

*corpus:* textDesc

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin

character data

*Note* Usually empty, unless some further clarification of the type attribute is needed, in which case it may contain running prose.

The list presented here is primarily for illustrative purposes.

*Example*

```
<domain type="domestic"/>
<domain type="rel">religious broadcast</domain>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element domain
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
```

```
   attribute type { text }?,
   macro.phraseSeq.limited}
```

---

**<edition>** (edition) describes the particularities of one edition of a text. [*2.2.2. The Edition Statement*]

*Module* header — The TEI Header

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.biblPart

*Contained by*

*core:* bibl monogr

*header:* editionStmt

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<edition>First edition <date>Oct 1990</date>
</edition>
<edition  n="S2">Students' edition</edition>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element edition
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

**<editionStmt>** (edition statement) groups information relating to one edition of a text. [*2.2.2. The Edition Statement 2.2. The File Description*]

*Module* header — The TEI Header

*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  - * *@ana*
- – att.global.change
  - * *@change*
- – att.global.facs
  - * *@facs*
- – att.global.linking
  - * *@corresp*
  - * *@synch*
  - * *@sameAs*
  - * *@copyOf*
  - * *@next*
  - * *@prev*
  - * *@exclude*
  - * *@select*
- – att.global.rendition
  - * *@rend*
  - * *@style*
  - * *@rendition*
- – att.global.responsibility
  - * *@cert*
  - * *@resp*
- – att.global.source
  - * *@source*

*Contained by* *header:* biblFull fileDesc

*May contain*

*core:* author editor meeting p respStmt

*header:* edition funder principal sponsor

*linking:* ab

*Example*

```
<editionStmt>
  <edition  n="S2">Students' edition</edition>
```

```
    <respStmt>
      <resp>Adapted by </resp>
      <name>Elizabeth Kirk</name>
    </respStmt>
  </editionStmt>
```

*Example*

```
  <editionStmt>
    <p>First edition, <date>Michaelmas Term, 1991.</date>
    </p>
  </editionStmt>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="edition"/>
   <classRef key="model.respLike"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element editionStmt
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.pLike+ | ( edition, model.respLike* ) )
}
```

---

**<editor>** contains a secondary statement of responsibility for a bibliographic item, for example the name of an individual, institution or organization, (or of several such) acting as editor, compiler, translator, etc. [*3.12.2.2. Titles, Authors, and Editors*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*

  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
  - att.naming
    - *@role*
    - *@nymRef*
  - att.canonical
    * *@key*
    * *@ref*

*Member of* model.respLike

*Contained by*

*core:* analytic bibl monogr series

*header:* editionStmt seriesStmt titleStmt

*msdescription:* msItem

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* A consistent format should be adopted.

Particularly where cataloguing is likely to be based on the content of the header, it is advisable to use generally recognized authority lists for the exact form of personal names.

*Example*

```
<editor role="Technical_Editor">Ron Van den Branden</editor>
<editor role="Editor-in-Chief">John Walsh</editor>
<editor role="Managing_Editor">Anne Baillot</editor>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element editor
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   macro.phraseSeq}
```

---

## <editorialDecl>

<editorialDecl> (editorial practice declaration) provides details of editorial principles and practices applied during the encoding of a text. [*2.3.3. The Editorial Practices Declaration 2.3. The Encoding Description 16.3.2. Declarable Elements*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.declarable
  - *@default*

*Member of* model.encodingDescPart
*Contained by*
*header:* encodingDesc
*May contain*
*core:* p
*header:* correction hyphenation interpretation normalization punctuation quotation segmentation stdVals
*linking:* ab
*Example*

```
<editorialDecl>
  <normalization>
    <p>All words converted to Modern American spelling using
        Websters 9th Collegiate dictionary
    </p>
  </normalization>
  <quotation  marks="all">
    <p>All opening quotation marks converted to “ all closing
        quotation marks converted to &cdq;.</p>
  </quotation>
</editorialDecl>
```

*Schematron* <sch:pattern is-a=”declarable”> <sch:param name=”tde” value=”tei:editorialDecl”/> </sch:pattern>
*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.pLike"/>
  <classRef key="model.editorialDeclPart"/>
 </alternate>
```

```
</content>
```

*Schema Declaration*

```
element editorialDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   ( model.pLike | model.editorialDeclPart )+
}
```

---

## <education>

(education) contains a description of the educational experience of a person. [*16.2.2. The Participant Description*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
    * *@key*
    * *@ref*
- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **primary**
**secondary**
**undergraduate**
**graduate**
**residency**
**apprenticeship**

*Member of* model.persStateLike
*Contained by*
*namesdates:* person personGrp persona
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit
*dictionaries:* lang oRef pRef
*figures:* figure formula notatedMusic
*gaiji:* g
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef specDesc specList tag val
*textcrit:* app witDetail
*textstructure:* floatingText
*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo
*verse:* caesura rhyme
character data
*Example*

```
<education>Left school at age 16</education>
```

*Example*

```
<education  from="1986-01-01"
  to="1990-06-30">Attended <name>Cherwell School</name>
</education>
```

*Example*

```
<education  notBefore="1685-07"
  notAfter="1690-06">Anthony Hammond
smuggled her into the University of Cambridge, where she was
disguised as his male cousin, Jack. She remained there for some
months learning grammar, logic, rhetoric, and ethics</education>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element education
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.phraseSeq}
```

---

**<eg>** (example) contains any kind of illustrative example. [*23.5. Element Specifications 23.5.3. Attribute List Specification*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.egLike

*Contained by*

*cmc:* post

*core:* add cit corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*tagdocs:* exemplum

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* If the example contains material in XML markup, either it must be enclosed within a CDATA marked section, or character entity references must be used to represent the markup delimiters. If the example contains well-formed XML, it should be marked using the more specific <egXML> element.

*Example*

```
<p>The
<gi>term</gi> element is declared using the following syntax:
<eg><![CDATA[<!ELEMENT term (%phrase.content;)>]]</eg>
</p>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element eg
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

**<egXML>** (example of XML) a single XML fragment demonstrating the use of some XML, such as elements, attributes, or processing instructions, etc., in which the <egXML> element functions as the root element. [*23.1.1. Phrase Level Terms*]

*Namespace* http://www.tei-c.org/ns/Examples

*Module* tagdocs — Documentation Elements

*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  * *@ana*
- att.global.change
  * *@change*
- att.global.facs
  * *@facs*
- att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

`@valid` indicates the intended validity of the example with respect to a schema.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **true** the example is intended to be fully valid, assuming that its root element, or a provided root element, could have been used as a possible root element in the schema concerned.[Default]

**feasible** the example could be transformed into a valid document by inserting any number of valid attributes and child elements anywhere within it; or it is valid

against a version of the schema concerned in which the provision of character data, list, element, or attribute values has been made optional.

**false** the example is not intended to be valid, and contains deliberate errors.

*Member of* model.egLike

*Contained by*

*cmc:* post

*core:* add cit corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*tagdocs:* exemplum

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain* ANY

*Note* In the source of the TEI Guidelines, this element declares itself and its content as belonging to the namespace `http://www.tei-c.org/ns/Examples`. This enables the content of the element to be validated independently against the TEI scheme. Where this element is used outside this context, a different namespace or none at all may be preferable. The content must however be a well-formed XML fragment or document: where this is not the case, the more general <eg> element should be used in preference.

*Example*

```
<egXML xmlns="http://www.tei-c.org/ns/Examples"><div>
   <head>A slide about <gi>egXML</gi>
   </head>
   <list>
     <item>
       <gi>egXML</gi> can be used to give XML examples in the TEI
           Examples namespace</item>
     <item>Attributes values for <att>valid</att>:
     <list  rend="collapsed">
        <item>
          <val  rend="green">true</val>: intended to be fully
                  valid</item>
        <item>
          <val  rend="amber">feasible</val>: valid if missing nodes
                  provided</item>
        <item>
          <val  rend="red">false</val>: not intended to be valid</item>
```

```
        </list>
      </item>
      <item>The <att>rend</att> attribute can be
            used for recording how parts of the example were rendered.</item>
   </list>
</div>
</egXML>
```

*Example*

```
<egXML  valid="feasible"  source="#UND"
xmlns="http://www.tei-c.org/ns/Examples"><text>
   <front>
<!-- front matter for the whole group -->
   </front>
   <group>
     <text>
<!-- first text -->
     </text>
     <text>
<!-- second text -->
     </text>
   </group>
</text>
<!-- This example is not valid TEI, but could be made so by
adding missing components -->
</egXML>
```

*Example*

```
<egXML xmlns="http://www.tei-c.org/ns/Examples" valid="false">
   <para xml:lang="en">Doubloons are a pirate's best friend</para>
</egXML>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <anyElement/>
 </alternate>
</content>
```

*Schema Declaration*

```
element egXML
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
```

```
   attribute valid { "true" | "feasible" | "false" }?,
   ( text | anyElement_egXML_2 )*
}
```

---

**<eLeaf>** (leaf or terminal node of an embedding tree) provides explicitly for a leaf of an embedding tree, which may also be encoded with the <eTree> element. [*20.3. Another Tree Notation*]

*Module* nets — Graphs, Networks, and Trees

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - *@type*
  - *@subtype*

`@value` indicates the value of an embedding leaf, which is a feature structure or other analytic element.

*Status* Optional
*Datatype* teidata.pointer

*Contained by:* eTree triangle
*May contain*
*core:* label ptr ref
*tagdocs:* listRef
*Note* The <eTree> tag may be used if the encoder does not wish to distinguish by name between nonleaf and leaf nodes in embedding trees; they are distinguished by their arrangement.
*Example*

```
<eLeaf value="http://an.fsurl.tei/#FSWITH">
  <label>with</label>
</eLeaf>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="label" minOccurs="0"/>
  <classRef key="model.ptrLike"
   minOccurs="0"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element eLeaf
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   attribute value { text }?,
   ( label?, model.ptrLike? )
}
```

---

**<elementRef>** points to the specification for some element which is to be included in a schema. [*23.2. Modules and Schemas*]

*Module* tagdocs — Documentation Elements
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.repeatable
  - *@minOccurs*
  - *@maxOccurs*

`@key` the identifier used for the required element within the source indicated.
*Status* Required
*Datatype* teidata.xmlName

*Member of* model.contentPart model.oddRef
*Contained by*
*tagdocs:* alternate content schemaSpec sequence specGrp
*May contain* Empty element
*Note* Elements are identified by the name supplied as value for the *ident* attribute on the <elementSpec> element in which they are declared. TEI element names are unique.
*Example*

```
<elementRef key="p"/>
```

Include in the schema an element named <p> available from the current default source.
*Example*

```
<elementRef  key="p"  source="tei:1.2.1"/>
```

Include in the schema an element named <p> available from the TEI P5 1.2.1 release.

*Schematron* <sch:pattern> <sch:rule context="tei:elementRef[ parent::tei:schemaSpec | parent::tei:specGrp ]"> <sch:report test="@minOccurs | @maxOccurs" role="error">An element reference is not repeatable when part of a schema specification (and thus this <elementRef> should not have @minOccurs or @maxOccurs).</sch:report> </sch:rule> <sch:rule context="tei:content//tei:elementRef"> <sch:report test="@source" role="error">An element reference within a content model must refer to a locally defined element specification (and thus this <elementRef> should not have @source).</sch:report> </sch:rule> </sch:pattern>

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element elementRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.repeatable.attributes,
   attribute key { text },
   empty
}
```

---

## <elementSpec>

(element specification) documents the structure, content, and purpose of a single element type. [*23.5. Element Specifications 23. Documentation Elements*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*

* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

• att.identified
– *@ident*
– *@predeclare*
– *@module*

– att.combinable
* *@mode*

* att.deprecated
· *@validUntil*

• att.namespaceable
– *@ns*

`@prefix` specifies a default prefix which will be prepended to all patterns relating to the element, unless otherwise stated.
*Status* Optional
*Datatype* 0–1 occurrences of teidata.xmlName separated by whitespace

*Member of* model.oddDecl

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*tagdocs:* schemaSpec specGrp

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc gloss

*tagdocs:* altIdent attList classes constraintSpec content equiv exemplum listRef model modelGrp modelSequence remarks valList

*Example*

```
<elementSpec module="tagdocs" ident="code">
  <gloss/>
  <desc>contains literal code</desc>
  <classes>
    <memberOf key="model.emphLike"/>
  </classes>
  <content>
    <textNode/>
  </content>
  <attList>
    <attDef ident="type" usage="opt">
      <desc>the language of the code</desc>
      <datatype>
        <dataRef key="teidata.enumerated"/>
      </datatype>
    </attDef>
  </attList>
</elementSpec>
```

*Schematron* <sch:rule context="tei:elementSpec[ @mode eq 'delete' ]"> <sch:report test="child::*">This elementSpec element has a mode= of "delete" even though it has child elements. Change the mode= to "add", "change", or "replace", or remove the child elements.</sch:report> </sch:rule> <sch:rule context="tei:elementSpec[ @mode = ('add','change','replace') ]"> <sch:assert test="child::* | (@* except (@mode, @ident))">This elementSpec element has a mode= of "<sch:value-of select="@mode"/>", but does not have any child elements or schema-changing attributes. Specify child elements, use validUntil=, predeclare=, ns=, or prefix=, or change the mode= to "delete".</sch:assert> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identSynonyms"/>
   <classRef key="model.descLike"/>
  </alternate>
  <elementRef key="classes" minOccurs="0"/>
  <elementRef key="content" minOccurs="0"/>
  <elementRef key="valList" minOccurs="0"/>
  <elementRef key="constraintSpec"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="attList" minOccurs="0"/>
```

```
  <alternate maxOccurs="unbounded"
   minOccurs="0">
   <elementRef key="model"/>
   <elementRef key="modelGrp"/>
   <elementRef key="modelSequence"/>
  </alternate>
  <elementRef key="exemplum" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="remarks" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="listRef" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element elementSpec
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.identified.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   att.namespaceable.attributes,
   attribute prefix { ? }?,
   (
      ( model.identSynonyms | model.descLike )*,
      classes?,
      content?,
      valList?,
      constraintSpec*,
      attList?,
      ( model | modelGrp | modelSequence )*,
      exemplum*,
      remarks*,
      listRef*
   )
}
```

---

**<ellipsis>** (deliberately marked omission) indicates a purposeful marking in the source document signalling that content has been omitted, and may also supply or describe the omitted content. [*3.5.3. Additions, Deletions, and Omissions*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*

  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.timed
  - *@start*

- *@end*

- att.duration
  - att.duration.iso
    - *@dur-iso*
  - att.duration.w3c
    - *@dur*

*Member of* model.global.edit

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*core:* desc

*transcr:* metamark supplied

*Note* Unlike <gap>, which indicates content that the encoder cannot or chooses not to represent, <ellipsis> indicates a passage explicitly signalled in the source document as absent. The <ellipsis> element is not appropriate for every use of ellipsis points, such as when they indicate that a speaker is pausing.

*Example*

```
<lg>
  <l>What projects men make—what queer turns they take,</l>
  <l>Since <emph>steam</emph> has improved our condition;</l>
  <l>They never are still, but must cure or must kill</l>
  <l>With steam physic or steam ammunition.</l>
  <l>But a short time ago, to a quack you would go,</l>
  <l>To steam a fat man to a thinner;</l>
  <l>Now changed from all that, if you wish to get <emph>fat</emph>,</l>
  <l>Come to Barton’s and eat a <emph>steam dinner!</emph>
  </l>
  <l>Oh dear! think of a scheme, odd though it seem—</l>
  <l>I’m sure ’twill succeed if you make it by steam.</l>
</lg>
<lg>
  <l>You may sleep, you may dream, you may travel by steam,</l>
  <l>For the outcry is still to go faster;</l>
  <l>And what does it reck, should you e’en break your neck,</l>
  <l>If ’tis <emph>steam</emph> that brings on the disaster?</l>
  <ellipsis  resp="#ChambersEdnbrghJrnl1880">
    <metamark  function="multilineEllipsis"> * * * * </metamark>
    <desc  resp="#teiProjectEditor2021">The printer omits four lines here,
        skipping the second half of the second octave, before the
refrain.</desc>
  </ellipsis>
  <l>Oh dear! think of a scheme, odd though it seem—</l>
  <l>I’m sure ’twill succeed if you make it by steam.</l>
</lg>
```

*Example*

```
<lg>
  <l>You think you’ve lost your love </l>
  <l>Well, I saw her yesterday </l>
  <l>It’s you she's thinking of </l>
  <l>And she told me what to say</l>
</lg>
<lg  xml:id="chorus">
  <label>[Refrain]</label>
  <l>She says she loves you </l>
  <l>And you know that can’t be bad </l>
  <l>Yes, she loves you </l>
  <l>And you know you should be glad</l>
</lg>
<lg>
  <l>She said you hurt her so </l>
  <l>She almost lost her mind </l>
  <l>But now she said she knows </l>
  <l>You’re not the hurting kind</l>
</lg>
<ellipsis>
  <metamark>******</metamark>
  <supplied  copyOf="#chorus"/>
</ellipsis>
```

*Content model*

```
<content>
 <sequence>
```

```
  <elementRef key="metamark"/>
  <classRef key="model.descLike"
   minOccurs="0" maxOccurs="1"/>
  <elementRef key="supplied" minOccurs="0"
   maxOccurs="1"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element ellipsis
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   ( metamark, model.descLike?, supplied? )
}
```

---

**<email>** (electronic mail address) contains an email address identifying a location to which email messages can be delivered. [*3.6.2. Addresses*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*

*Member of* model.addressLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The format of a modern Internet email address is defined in RFC 2822

*Example*

```
<email>membership@tei-c.org</email>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element email
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   macro.phraseSeq}
```

**<emph>** (emphasized) marks words or phrases which are stressed or emphasized for linguistic or rhetorical effect. [*3.3.2.2. Emphatic Words and Phrases 3.3.2. Emphasis, Foreign Words, and Unusual Language*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.written
  - *@hand*

*Member of* model.emphLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
You took the car and did <emph>what</emph>?!!
```

*Example*

```
<q>What it all comes to is this,</q> he said.
<q>
  <emph>What
     does Christopher Robin do in the morning nowadays?</emph>
</q>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element emph
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.written.attributes,
   macro.paraContent}
```

---

**<empty>** indicates the presence of an empty node within a content model. [*23. Documentation Elements*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.contentPart
*Contained by*
*tagdocs:* alternate content sequence
*May contain* Empty element
*Example*

```
<content>
  <alternate  minOccurs="0"
    maxOccurs="unbounded">
    <empty/>
    <elementRef  key="hi"/>
  </alternate>
</content>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element empty
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
```

```
   att.global.responsibility.attributes,
   att.global.source.attributes,
   empty
}
```

---

## <encodingDesc>

<encodingDesc> (encoding description) documents the relationship between an electronic text and the source or sources from which it was derived. [*2.3. The Encoding Description 2.1.1. The TEI Header and Its Components*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Member of* model.teiHeaderPart

*Contained by*

*header:* teiHeader

*May contain*

*core:* p

*gaiji:* charDecl

*header:* appInfo classDecl editorialDecl geoDecl listPrefixDef projectDesc refsDecl samplingDecl schemaRef styleDefDecl tagsDecl unitDecl

*iso-fs:* fsdDecl

*linking:* ab

*spoken:* transcriptionDesc

*tagdocs:* constraintDecl schemaSpec

*textcrit:* variantEncoding

*verse:* metDecl

*Example*

```
<encodingDesc>
  <p>Basic encoding, capturing lexical information only. All
     hyphenation, punctuation, and variant spellings normalized. No
     formatting or layout information preserved.</p>
</encodingDesc>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.encodingDescPart"/>
  <classRef key="model.pLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element encodingDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.encodingDescPart | model.pLike )+
}
```

---

**<entry>** (entry) contains a single structured entry in any kind of lexical resource, such as a dictionary or lexicon. [*10.1. Dictionary Body and Overall Structure 10.2. The Structure of Dictionary Entries*]

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.entryLike
  - *@type*

- att.typed
  - ~~type~~
  - @subtype

- att.sortable
  - *@sortKey*

*Member of* model.entryLike model.entryPart.top
*Contained by*
*dictionaries:* entry hom re sense superEntry
*drama:* castList epilogue performance prologue set
*figures:* figure
*linking:* standOff
*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript
*May contain*
*analysis:* interp interpGrp pc span spanGrp

*certainty:* certainty precision respons
*core:* bibl biblStruct cb cit ellipsis gap gb index lb listBibl milestone note noteGrp pb ptr ref
*dictionaries:* def dictScrap entry etym form gramGrp hom lbl re sense usg xr
*figures:* figure notatedMusic
*header:* biblFull
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* msDesc
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* listRef
*textcrit:* app witDetail
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* Like all elements, <entry> inherits an *xml:id* attribute from the class *global*. No restrictions are placed on the method used to construct *xml:id*s; one convenient method is to use the orthographic form of the headword, appending a disambiguating number where necessary. Identification codes are sometimes included on machine-readable tapes of dictionaries for in-house use.

It is recommended to use the <sense> element even for an entry that has only one sense to group together all parts of the definition relating to the word sense since this leads to more consistent encoding across entries.

*Example*

```
<entry>
  <form>
    <orth>disproof</orth>
    <pron>dIs"pru:f</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
  <sense  n="1">
    <def>facts that disprove something.</def>
  </sense>
  <sense  n="2">
    <def>the act of disproving.</def>
  </sense>
</entry>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <elementRef key="hom"/>
  <elementRef key="sense"/>
  <elementRef key="pc"/>
  <classRef key="model.entryPart.top"/>
  <classRef key="model.global"/>
  <classRef key="model.ptrLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element entry
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.entryLike.attributes,
   att.sortable.attributes,
   ( hom | sense | pc | model.entryPart.top | model.global | model.ptrLike )+
}
```

---

**<entryFree>** (unstructured entry) contains a single unstructured entry in any kind of lexical resource, such as a dictionary or lexicon. [*10.1. Dictionary Body and Overall Structure 10.2. The Structure of Dictionary Entries*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

    * *@cert*
    * *@resp*
  – att.global.source
    * *@source*
- att.entryLike
  – *@type*
  – att.typed
    * ~~type~~
    * @subtype
- att.lexicographic
  – *@expand*
  – *@split*
  – *@value*
  – *@location*
  – *@mergedIn*
  – *@opt*
  – att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  – att.lexicographic.normalized
    * *@norm*
    * *@orig*
- att.sortable
  – *@sortKey*

*Member of* model.entryLike

*Contained by*

*drama:* castList epilogue performance prologue set

*figures:* figure

*linking:* standOff

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index label lb list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* case colloc def etym form gen gram gramGrp hom hyph iType lang lbl mood number oRef orth pRef per pos pron re sense subc superEntry syll tns usg xr

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May contain any dictionary elements in any combination.

*Example*

```
<entryFree>
  <orth>biryani</orth> or <orth>biriani</orth>
  <pron>(%bIrI"A:nI)</pron>
  <def>any of a variety of Indian dishes ...</def>
  <etym>[from <lang>Urdu</lang>]</etym>
</entryFree>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.entryPart"/>
  <classRef key="model.morphLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element entryFree
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.entryLike.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.sortable.attributes,
   (
      text
    | model.gLike    | model.entryPart    | model.morphLike    | model.phrase    | model.i
}
```

---

**<epigraph>** (epigraph) contains a quotation, anonymous or attributed, appearing at the start or end of a section or on a title page. [*4.2.3. Arguments, Epigraphs, and Postscripts 4.2. Elements Common to All Divisions 4.6. Title Pages*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

    * *@source*

  - att.cmc
    - *@generatedBy*

*Member of* model.divWrapper model.pLike.front model.titlepagePart

*Contained by*

*cmc:* post

*core:* lg list

*drama:* castList epilogue performance prologue

*figures:* figure table

*msdescription:* msItem

*textcrit:* lem rdg

*textstructure:* back body div div1 div2 div3 div4 div5 div6 div7 front group opener titlePage

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct cb cit desc ellipsis gap gb index l label lb lg list listBibl milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Example*

```
<epigraph  xml:lang="la">
  <cit>
    <bibl>Lucret.</bibl>
    <quote>
      <l  part="F">petere inde coronam,</l>
      <l>Vnde prius nulli velarint tempora Musae.</l>
    </quote>
  </cit>
</epigraph>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
```

```
  <classRef key="model.common"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element epigraph
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   ( model.common | model.global )*
}
```

---

**<epilogue>** (epilogue) contains the epilogue to a drama, typically spoken by an actor out of character, possibly in association with a particular performance or venue. [*7.1.2. Prologues and Epilogues 7.1. Front and Back Matter* ]

*Module* drama — Performance Texts

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.frontPart.drama

*Contained by*

*textcrit:* lem rdg

*textstructure:* back front

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct cb cit desc ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* Contains optional headings, a sequence of one or more component-level elements, and an optional sequence of closing material.

*Example*

```
<epilogue>
  <head>Written by <name>Colley Cibber, Esq</name> and spoken by <name>Mrs.
Cibber</name>
  </head>
  <sp>
    <lg  type="couplet">
      <l>Since Fate has robb'd me of the hapless Youth,</l>
```

```
      <l>For whom my heart had hoarded up its truth;</l>
    </lg>
    <lg  type="couplet">
      <l>By all the Laws of Love and Honour, now,</l>
      <l>I'm free again to chuse, — and one of you</l>
    </lg>
    <lg  type="triplet">
      <l>Suppose I search the sober Gallery; — No,</l>
      <l>There's none but Prentices — & Cuckolds all a row:</l>
      <l>And these, I doubt, are those that make 'em so.</l>
    </lg>
    <stage  type="business">Pointing to the Boxes.</stage>
    <lg  type="couplet">
      <l>'Tis very well, enjoy the jest:</l>
    </lg>
  </sp>
</epilogue>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.common"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divBottom"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element epilogue
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      ( model.divTop | model.global )*,
      ( ( model.common, model.global* )+ ),
```

```
      ( ( model.divBottom, model.global* )* )
   )
}
```

---

**<equipment>** (equipment) provides technical details of the equipment and media used for an audio or video recording used as the source for a spoken text. [*8.2. Documenting the Source of Transcribed Speech 16.3.2. Declarable Elements*]

*Module* spoken — Transcriptions of Speech

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable
  - *@default*

*Member of* model.recordingPart

*Contained by*
*spoken:* recording
*May contain*
*core:* p
*linking:* ab
*Example*

```
<equipment>
  <p>"Hi-8" 8 mm NTSC camcorder with integral directional
     microphone and windshield and stereo digital sound
     recording channel.
  </p>
</equipment>
```

*Example*

```
<equipment>
  <p>8-track analogue transfer mixed down to 19 cm/sec audio
     tape for cassette mastering</p>
</equipment>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:equipment"/> </sch:pattern>
*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element equipment
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   model.pLike+
}
```

---

**<equiv>** (equivalent) specifies a component which is considered equivalent to the parent element, either by co-reference, or by external link. [*3.4.1. Terms and Glosses 23.4.1. Description of Components*]

*Module* tagdocs — Documentation Elements
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.internetMedia
  - *@mimeType*
- att.predicate
  - *@predicate*

`@name` a single word which follows the rules defining a legal XML name (see `https://www.w3.org/TR/REC-xml/#dt-name`), naming the underlying concept of which the parent is a representation.
*Status* Optional
*Datatype* teidata.name

`@uri` (uniform resource identifier) references the underlying concept of which the parent is a representation by means of some external identifier.
*Status* Optional
*Datatype* teidata.pointer

`@filter` references an external script which contains a method to transform instances of this element to canonical TEI.
*Status* Optional
*Datatype* teidata.pointer

*Member of* model.identEquiv

*Contained by*

*header:* category taxonomy

*linking:* joinGrp

*tagdocs:* attDef classSpec constraintDecl constraintSpec dataSpec elementSpec macroSpec model modelGrp modelSequence moduleSpec paramSpec schemaSpec valItem

*May contain* Empty element

*Note* The *mimeType* attribute should be used to supply the MIME media type of the filter script specified by the *filter* attribute.

*Example* The following example declares that the <bo> element is conceptually equivalent to the markup construct <hi rend='bold'>, and that an external definition of this concept is available from the URI indicated

```
<elementSpec ident="hi" mode="change">
  <equiv name="BOLD"/>
  <desc>bold typography</desc>
  <attList>
    <attDef ident="rend" mode="change">
      <valList>
        <valItem ident="bold"/>
      </valList>
    </attDef>
  </attList>
</elementSpec>
<elementSpec ident="bo" mode="add">
  <equiv name="BOLD"
    uri="http://www.example.com/typesetting/bold"/>
</elementSpec>
```

*Content model* <content> <empty/></content>

*Schema Declaration*

```
element equiv
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.internetMedia.attributes,
   att.predicate.attributes,
   attribute name { text }?,
   attribute uri { text }?,
   attribute filter { text }?,
   empty
}
```

---

## <eTree>

<eTree> (embedding tree) provides an alternative to the <tree> element for representing ordered rooted tree structures. [*20.3. Another Tree Notation*]

*Module* nets — Graphs, Networks, and Trees

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.typed
  - *@type*
  - *@subtype*

`@value` provides the value of an embedding tree, which is a feature structure or other analytic element.
*Status* Optional
*Datatype* teidata.pointer

*Member of* model.divPart model.standOffPart
*Contained by*
*core:* item note q quote said stage textLang
*drama:* castList epilogue performance prologue set view
*figures:* cell figure
*header:* change handNote licence scriptNote
*linking:* standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*nets:* eTree forest triangle

*tagdocs:* specGrp

*textcrit:* lem rdg

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript

*transcr:* metamark

*May contain*

*core:* label ptr ref

*nets:* eLeaf eTree triangle

*tagdocs:* listRef

*Note* an optional label followed by zero or more embedding trees, triangles, or embedding leafs.

*Example*

```
<eTree  n="ex1">
  <label>PP</label>
  <eTree>
    <label>P</label>
    <eLeaf>
      <label>with</label>
    </eLeaf>
  </eTree>
  <eTree>
    <label>NP</label>
    <eTree>
      <label>Art</label>
      <eLeaf>
        <label>the</label>
      </eLeaf>
    </eTree>
    <eTree>
      <label>N</label>
      <eLeaf>
        <label>periscope</label>
      </eLeaf>
    </eTree>
  </eTree>
</eTree>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="label" minOccurs="0"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="eTree"/>
   <elementRef key="triangle"/>
   <elementRef key="eLeaf"/>
   <classRef key="model.ptrLike"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element eTree
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   attribute value { text }?,
   ( label?, ( eTree | triangle | eLeaf | model.ptrLike )* )
}
```

---

**<etym>** (etymology) encloses the etymological information in a dictionary entry. [*10.3.4. Etymological Information*]

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*

- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*
- att.lexicographic.normalized
  - *@norm*
  - *@orig*

- att.typed
  - *@type*
  - *@subtype*

*Member of* model.entryPart model.entryPart.top

*Contained by*

*core:* cit

*dictionaries:* dictScrap entry entryFree etym hom re sense

*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index label lb list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* def etym gramGrp lang lbl oRef pRef usg xr

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat

geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May contain character data mixed with any other elements defined in the dictionary tag set.

There is no consensus on the internal structure of etymologies, or even on whether such a structure is appropriate. The <etym> element accordingly simply contains prose, within which names of languages, cited words, or parts of words, glosses, and examples will typically be prominent. The tagging of such internal objects is optional.

*Example*

```
<entry>
  <form>
    <orth>publish</orth> ... </form>
  <etym>
    <lang>ME.</lang>
    <mentioned>publisshen</mentioned>,
  <lang>F.</lang>
    <mentioned>publier</mentioned>, <lang>L.</lang>
    <mentioned>publicare,
        publicatum</mentioned>. <xr>See <ref>public</ref>; cf. 2d
<ref>-ish</ref>.</xr>
  </etym>
</entry> (From: Webster's Second International)
```

*Example*

```
<entry>
  <form>
    <orth>Handschuh</orth> ... </form>
  <etym  type="compounding">
    <oRef>Hand</oRef> (<pRef  notation="ipa">'hant</pRef>): <gloss>hand</gloss>,
  <etym  type="metaphor">
     <oRef>Schuh</oRef> (<pRef  notation="ipa">ʃu:</pRef>): <gloss>shoe</gloss>
    </etym>
  </etym>
</entry>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.global"/>
  <classRef key="model.inter"/>
  <classRef key="model.phrase"/>
```

```
  <elementRef key="def"/>
  <elementRef key="etym"/>
  <elementRef key="gramGrp"/>
  <elementRef key="lbl"/>
  <elementRef key="usg"/>
  <elementRef key="xr"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element etym
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.typed.attributes,
   (
      text
    | model.gLike    | model.global    | model.inter    | model.phrase    | def    | etym
}
```

---

**<event>** (event) contains data relating to anything of significance that happens in time. [*14.3.1. Basic Principles*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.locatable
  - *@where*
- att.naming

- *@role*
- *@nymRef*

- att.canonical
  - *@key*
  - *@ref*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.eventLike

*Contained by*

*namesdates:* event listEvent org person personGrp persona place

*May contain*

*core:* bibl biblStruct desc head label listBibl note noteGrp p ptr ref

*header:* biblFull idno

*linking:* ab link linkGrp

*msdescription:* msDesc

*namesdates:* event eventName listEvent listObject listPerson listPlace listRelation object org person personGrp place relation

*tagdocs:* listRef

*Example*

```
<listEvent>
  <event when="1618-05-23"
    xml:id="SecondDefPrague" where="#Prague">
    <eventName>1618 Defenestration of Prague</eventName>
    <idno>https://www.wikidata.org/wiki/Q13365740</idno>
    <listPerson type="defenstrated">
      <person>
        <persName>Jaroslav Bořita z Martinic</persName>
        <idno type="GND">https://d-nb.info/gnd/116810998</idno>
      </person>
      <person>
        <persName>Vilém Slavata z Chlumu a Košumberka</persName>
        <idno type="GND">https://d-nb.info/gnd/1018376615</idno>
      </person>
      <person>
        <persName>Filip Fabricius</persName>
        <idno type="GND">https://d-nb.info/gnd/133946118</idno>
      </person>
    </listPerson>
    <place xml:id="Prague">
      <placeName>Prague</placeName>
    </place>
  </event>
  <event from="1618" to="1648"
    xml:id="ThirtyYearsWar">
    <eventName>Thirty Years' War</eventName>
    <idno>https://www.wikidata.org/wiki/Q2487</idno>
```

```
      <event  when="1643-03-19"
        xml:id="BattleofRocroi"  where="#Rocroi">
        <eventName>Battle of Rocroi</eventName>
        <idno  type="Wikidata">https://www.wikidata.org/wiki/Q728480</idno>
        <idno  type="GND">https://d-nb.info/gnd/4202901-6</idno>
        <place  xml:id="Rocroi">
          <placeName>Rocroi</placeName>
          <location>
            <geo  decls="#WGS">49.926111 4.522222</geo>
          </location>
        </place>
      </event>
    </event>
  </listEvent>
```

*Example*

```
  <person>
    <event  type="mat"  when="1972-10-12">
      <label>matriculation</label>
    </event>
    <event  type="grad"  when="1975-06-23">
      <label>graduation</label>
    </event>
  </person>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="idno" minOccurs="0"
   maxOccurs="unbounded"/>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate>
   <classRef key="model.pLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <classRef key="model.labelLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <elementRef key="eventName"
    minOccurs="1" maxOccurs="unbounded"/>
  </alternate>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.noteLike"/>
   <classRef key="model.biblLike"/>
   <classRef key="model.ptrLike"/>
   <elementRef key="linkGrp"/>
   <elementRef key="link"/>
   <elementRef key="idno"/>
  </alternate>
  <classRef key="model.eventLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.personLike"
    minOccurs="1" maxOccurs="1"/>
   <elementRef key="listPerson"
```

```
    minOccurs="1" maxOccurs="1"/>
  </alternate>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.placeLike"
    minOccurs="1" maxOccurs="1"/>
   <elementRef key="listPlace"
    minOccurs="1" maxOccurs="1"/>
  </alternate>
  <classRef key="model.objectLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="relation" minOccurs="1"
    maxOccurs="1"/>
   <elementRef key="listRelation"
    minOccurs="1" maxOccurs="1"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element event
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.locatable.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      idno*,
      model.headLike*,
      ( model.pLike+ | model.labelLike+ | eventName+ ),
      (
         model.noteLike       | model.biblLike       | model.ptrLike       | linkGrp
      model.eventLike*,
      ( model.personLike | listPerson )*,
      ( model.placeLike | listPlace )*,
      model.objectLike*,
      ( relation | listRelation )*
   )
}
```

---

**<eventName>** (name of an event) contains a proper noun or noun phrase used to refer to an event. [*14.2.4. Event Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*

    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  – att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  – att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  – *@evidence*
  – *@instant*
- att.personal
  – *@full*
  – *@sort*
  – att.naming
    * *@role*
    * *@nymRef*
    * att.canonical
      · *@key*
      · *@ref*
- att.typed
  – *@type*
  – *@subtype*

*Member of* model.nameLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education event eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<listEvent>
  <event from="1939-09-01" to="1945-09-02">
    <eventName xml:lang="de">Zweiter Weltkrieg</eventName>
    <eventName xml:lang="en">World War II</eventName>
    <idno type="GND">https://d-nb.info/gnd/4079167-1</idno>
    <idno type="Wikidata">https://www.wikidata.org/wiki/Q362</idno>
    <event from="1939-09-01" to="1939-10-06"
      xml:id="UeberfallAufPolen">
      <eventName xml:lang="de">Überfall auf Polen</eventName>
      <eventName xml:lang="en">Invasion of Poland</eventName>
      <idno type="GND">https://d-nb.info/gnd/4175002-0</idno>
      <idno type="LOC">https://id.loc.gov/authorities/sh85148341</idno>
      <listPlace type="affected">
        <place>
          <placeName xml:lang="pl">Gdańsk</placeName>
          <location>
            <geo>54.350556 18.652778</geo>
          </location>
        </place>
      </listPlace>
    </event>
    <event from="1941-06-22" to="1945-05-09">
      <eventName xml:lang="de">Deutsch-Sowjetischer Krieg</eventName>
      <eventName xml:lang="ru">Великая Отечественная война</eventName>
      <idno type="GND">https://d-nb.info/gnd/4076906-9</idno>
      <idno type="Wikidata">https://www.wikidata.org/wiki/Q189266</idno>
    </event>
  </event>
</listEvent>
```

*Example*

```
<p>On <date when="1719-03-19">Monday</date>, <rs type="person">she</rs> was
writing about the
<eventName ref="#SecondDefPrague">1618 Defenestration of Prague</eventName>
which initiated the
<rs type="event" ref="#ThirtyYearsWar">long war</rs>.</p>
```

*Example*

```
<event from="2019-09-16" to="2019-09-20"
  xml:id="tei2019graz">
  <eventName type="full">TEI 2019: What is text, really? TEI and
beyond</eventName>
  <eventName type="short">TEI 2019</eventName>
  <note> The abstract leading to the <gi>eventName</gi> element is available at
<ref target="https://gams.uni-graz.at/o:tei2019.141">https://gams.uni-graz.at/o:tei2019.141</
    Other related documents are available through
<ref target="https://gams.uni-graz.at/tei2019">https://gams.uni-graz.at/tei2019</ref>,
as well as in the
  <ref target="https://zenodo.org/communities/tei2019">TEI 2019 Zenodo
community</ref>.
  </note>
  <listPerson type="LocalOrganizers">
    <person>
      <persName>
        <surname>Raunig</surname>
```

```
          <forename>Elisabeth</forename>
        </persName>
      </person>
      <person>
        <persName>
          <surname>Scholger</surname>
          <forename>Martina</forename>
        </persName>
      </person>
      <person>
        <persName>
          <surname>Scholger</surname>
          <forename>Walter</forename>
        </persName>
      </person>
      <person>
        <persName>
          <surname>Steiner</surname>
          <forename>Elisabeth</forename>
        </persName>
      </person>
      <person>
        <persName>
          <surname>Vogeler</surname>
          <forename>Georg</forename>
        </persName>
      </person>
    </listPerson>
    <place  xml:lang="de">
      <placeName>Universität Graz</placeName>
      <location>
        <address>
          <addrLine>ReSoWi Gebäude</addrLine>
          <addrLine>Universitätsstraße 15</addrLine>
          <postCode>8010</postCode>
          <settlement>Graz</settlement>
          <country>Österreich</country>
        </address>
        <geo>15.451651587656 47.078215112534</geo>
      </location>
    </place>
    <listRelation>
      <relation  active="#tei2019graz"
        passive="#AnnualTEIConference"  type="CRM"  name="P31_is_instance_of"
        ref="https://www.wikidata.org/wiki/Property:P31"/>
    </listRelation>
  </event>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element eventName
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.personal.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<ex>** (editorial expansion) contains a sequence of letters added by an editor or transcriber when expanding an abbreviation. [*12.3.1.2. Abbreviation and Expansion*]

*Module* transcr — Representation of Primary Sources

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.editLike
  - *@evidence*
  - *@instant*

*Member of* model.choicePart model.pPart.editorial

*Contained by*

*analysis:* cl pc phr s span w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*gaiji:* g

character data

*Example*

```
The address is Southmoor <choice>
  <expan>R<ex>oa</ex>d</expan>
  <abbr>Rd</abbr>
</choice>
```

*Content model* `<content> <macroRef key="macro.xtext"/></content>`

*Schema Declaration*

```
element ex
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   macro.xtext}
```

---

## <exemplum>

(exemplum) groups an example demonstrating the use of an element along with optional paragraphs of commentary. [*23.5. Element Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

  * *@ana*
  - att.global.change
  * *@change*
  - att.global.facs
  * *@facs*
  - att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
  - att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
  - att.global.responsibility
  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.translatable
  - *@versionDate*
- att.typed
  - *@type*
  - *@subtype*

*Contained by* *tagdocs:* attDef classSpec dataSpec elementSpec macroSpec moduleSpec
*May contain*
*core:* p
*linking:* ab
*tagdocs:* eg egXML
*Example*

```
<exemplum  xml:lang="en">
  <p>The <gi>name</gi> element can be used for both personal names and place
names:</p>
  <eg  xml:space="preserve"><![CDATA[ <q>My dear <name type="person">Mr.
          Bennet</name>,</q> said his lady to him one day,
          <q>have you heard that <name type="place">Netherfield
          Park</name> is let at last?</q>]]></eg>
  <p>As shown above, the <att>type</att> attribute may be used to distinguish
the one from the
```

```
        other.</p>
    </exemplum>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.pLike" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate>
   <elementRef key="egXML"/>
   <elementRef key="eg"/>
  </alternate>
  <classRef key="model.pLike" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element exemplum
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.translatable.attributes,
   att.typed.attributes,
   ( model.pLike*, ( egXML | eg ), model.pLike* )
}
```

---

**<expan>** (expansion) contains the expansion of an abbreviation. [*3.6.5. Abbreviations and Their Expansions*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*

  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.editLike
  - *@evidence*
  - *@instant*

*Member of* model.choicePart model.pPart.editorial

*Contained by*

*analysis:* cl pc phr s span w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName

occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The content of this element should be the expanded abbreviation, usually (but not always) a complete word or phrase. The <ex> element provided by the `transcr` module may be used to mark up sequences of letters supplied within such an expansion.

If abbreviations are expanded silently, this practice should be documented in the <editorialDecl>, either with a <normalization> element or a <p>.

*Example*

```
The address is Southmoor
<choice>
  <expan>Road</expan>
  <abbr>Rd</abbr>
</choice>
```

*Example*

```
<choice  xml:lang="la">
  <abbr>Imp</abbr>
  <expan>Imp<ex>erator</ex>
  </expan>
</choice>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element expan
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.editLike.attributes,
   macro.phraseSeq}
```

---

**<explicit>** (explicit) contains the *explicit* of a item, that is, the closing words of the text proper, exclusive of any rubric or colophon which might follow it. [*11.6.1. The msItem and msItemStruct Elements*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.msExcerpt
  - *@defective*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.biblPart model.msQuoteLike

*Contained by*

*core:* bibl

*msdescription:* msItem msItemStruct

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme
character data

*Example*

```
<explicit>sed libera nos a malo.</explicit>
<rubric>Hic explicit oratio qui dicitur dominica.</rubric>
<explicit  type="defective">ex materia quasi et forma sibi
proporti<gap/>
</explicit>
<explicit  type="reverse">saued be shulle that doome of day the at
</explicit>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element explicit
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.msExcerpt.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<extent>** (extent) describes the approximate size of a text stored on some carrier medium or of some other object, digital or non-digital, specified in any convenient units. [*2.2.3. Type and Extent of File 2.2. The File Description 3.12.2.4. Imprint, Size of a Document, and Reprint Information 11.7.1. Object Description*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs

    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.biblPart

*Contained by*

*core:* bibl monogr

*header:* biblFull fileDesc

*msdescription:* supportDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<extent>3200 sentences</extent>
<extent>between 10 and 20 Mb</extent>
<extent>ten 3.5 inch high density diskettes</extent>
```

*Example* The <measure> element may be used to supply normalized or machine tractable versions of the size or sizes concerned.

```
<extent>
  <measure  unit="MiB"  quantity="4.2">About four megabytes</measure>
  <measure  unit="pages"  quantity="245">245 pages of source
     material</measure>
</extent>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element extent
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

---

**<f>** (feature) represents a *feature value specification*, that is, the association of a name with a value of any of several different types. [*19.2. Elementary Feature Structures and the Binary Feature Value*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*

`@name` a single word which follows the rules defining a legal XML name (see `https://www.w3.org/TR/REC-xml/#dt-name`), providing a name for the feature.
*Status* Required
*Datatype* teidata.name

`@fVal` (feature value) references any element which can be used to represent the value of a feature.
*Status* Optional
*Datatype* teidata.pointer
*Note* If this attribute is supplied as well as content, the value referenced is to be unified with that contained.

*Contained by:* ~~iso-fs~~ bicond cond fLib fs if
*May contain*
*iso-fs:* binary default fs numeric string symbol vAlt vColl vLabel vMerge vNot
character data
*Note* The content of <f> may be textual, with the assumption that the data type of the feature value is determined by the schema—this is the approach used in many language-technology-oriented projects and recommendations.
*Example*

```
<f  name="gender">
  <symbol  value="feminine"/>
</f>
```

*Example*

```
<fs>
  <f  name="voice">active</f>
  <f  name="tense">SimPre</f>
</fs>
```

*Content model*

```
<content>
 <alternate minOccurs="1" maxOccurs="1">
  <textNode/>
  <classRef key="model.featureVal"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element f
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datcat.attributes,
   attribute name { text },
   attribute fVal { text }?,
   ( text | model.featureVal )
}
```

---

**<facsimile>** contains a representation of some written source in the form of a set of images rather than as transcribed or encoded text. [*12.1. Digital Facsimiles*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.declaring
  - *@decls*

*Member of* model.resource
*Contained by*
*core:* teiCorpus
*textstructure:* TEI
*transcr:* facsimile
*May contain*
*core:* binaryObject graphic media
*figures:* formula
*textstructure:* back front
*transcr:* facsimile surface surfaceGrp
*Example*

```
<facsimile>
  <graphic url="page1.png"/>
  <surface>
    <graphic url="page2-highRes.png"/>
    <graphic url="page2-lowRes.png"/>
  </surface>
  <graphic url="page3.png"/>
  <graphic url="page4.png"/>
</facsimile>
```

*Example*

```
<facsimile>
  <surface ulx="0" uly="0" lrx="200" lry="300">
    <graphic url="Bovelles-49r.png"/>
  </surface>
</facsimile>
```

*Schematron* <sch:rule context="tei:facsimile//tei:line | tei:facsimile//tei:zone"> <sch:report test="child::text()[ normalize-space(.) ne "]"> A facsimile element represents a text with images, thus transcribed text should not be present within it. </sch:report> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <elementRef key="front" minOccurs="0"/>
  <alternate>
   <alternate minOccurs="1"
    maxOccurs="unbounded">
    <classRef key="model.graphicLike"/>
    <elementRef key="surface"/>
    <elementRef key="surfaceGrp"/>
   </alternate>
   <elementRef key="facsimile"
    minOccurs="1" maxOccurs="unbounded"/>
  </alternate>
  <elementRef key="back" minOccurs="0"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element facsimile
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   (
      front?,
      ( ( model.graphicLike | surface | surfaceGrp )+ | facsimile+ ),
      back?
   )
}
```

---

## <factuality>

(factuality) describes the extent to which the text may be regarded as imaginative or non-imaginative, that is, as describing a fictional or a non-fictional world. [*16.2.1. The Text Description*]

*Module* corpus — Language Corpora

*Attributes*

- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.typed
  - ~~type~~
  - @subtype

`@type` categorizes the factuality of the text.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **fiction** the text is to be regarded as entirely imaginative

**fact** the text is to be regarded as entirely informative or factual

**mixed** the text contains a mixture of fact and fiction

**inapplicable** the fiction/fact distinction is not regarded as helpful or appropriate to this text

*Member of* model.textDescPart

*Contained by*

*corpus:* textDesc

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin

character data

*Note* Usually empty, unless some further clarification of the type attribute is needed, in which case it may contain running prose

For many literary texts, a simple binary opposition between 'fiction' and 'fact' is naïve in the extreme; this parameter is not intended for purposes of subtle literary analysis, but as a simple means of characterizing the claimed fictiveness of a given text. No claim is made that works characterized as 'fact' are in any sense 'true'.

*Example*

```
<factuality type="fiction"/>
```

*Example*

```
<factuality type="mixed">contains a mixture of gossip and
speculation about real people and events</factuality>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element factuality
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { "fiction" | "fact" | "mixed" | "inapplicable" }?,
   macro.phraseSeq.limited}
```

---

**<faith>** (faith) specifies the faith, religion, or belief set of a person. [*14.3.2.1. Personal Characteristics*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.datable

- *@period*

- att.datable.custom
  - *@when-custom*
  - *@notBefore-custom*
  - *@notAfter-custom*
  - *@from-custom*
  - *@to-custom*
  - *@datingPoint*
  - *@datingMethod*
- att.datable.iso
  - *@when-iso*
  - *@notBefore-iso*
  - *@notAfter-iso*
  - *@from-iso*
  - *@to-iso*
- att.datable.w3c
  - *@when*
  - *@notBefore*
  - *@notAfter*
  - *@from*
  - *@to*

- att.editLike
  - *@evidence*
  - *@instant*
- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **practicing**
**clandestine**
**patrilineal**
**matrilineal**
**convert**

*Member of* model.persStateLike
*Contained by*
*namesdates:* person personGrp persona
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<faith>protestant</faith>
```

*Example*

```
<faith  ref="http://dbpedia.org/page/Manichaeism">Manichaeism</faith>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element faith
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
```

```
   att.editLike.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.phraseSeq}
```

---

**<fDecl>** (feature declaration) declares a single feature, specifying its name, organization, range of allowed values, and optionally its default value. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*

`@name` a single word which follows the rules defining a legal XML name (see `https://www.w3.org/TR/REC-xml/#dt-name`), indicating the name of the feature being declared; matches the *name* attribute of <f> elements in the text.
*Status* Required
*Datatype* teidata.name

`@optional` indicates whether or not the value of this feature may be present.
*Status* Optional
*Datatype* teidata.truthValue
*Default* true
*Note* If a feature is marked as optional, it is possible for it to be omitted from a feature structure. If an obligatory feature is omitted, then it is understood to have a default value, either explicitly declared, or, if no default is supplied, the special value any. If an optional feature is omitted, then it is understood to be missing and any possible value (including the default) is ignored.

*Contained by:* *iso-fs:* fsDecl
*May contain*
*iso-fs:* fDescr vDefault vRange
*Example*

```
<fDecl name="INV">
  <fDescr>inverted sentence</fDescr>
  <vRange>
    <vAlt>
      <binary value="true"/>
      <binary value="false"/>
    </vAlt>
  </vRange>
  <vDefault>
    <binary value="false"/>
  </vDefault>
</fDecl>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="fDescr" minOccurs="0"/>
  <elementRef key="vRange"/>
  <elementRef key="vDefault" minOccurs="0"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element fDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.datcat.attributes,
   attribute name { text },
   attribute optional { text }?,
   ( fDescr?, vRange, vDefault? )
}
```

---

**<fDescr>** (feature description (in FSD)) describes in prose what is represented by the feature being declared and its values. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by:* iso-fs: fDecl

*May contain*

*core:* abbr address bibl biblStruct choice cit date desc distinct email emph expan foreign gloss hi label list listBibl measure measureGrp mentioned name num ptr q quote ref rs said soCalled stage term time title unit

*dictionaries:* lang

*drama:* camera caption castList move sound tech view

*figures:* table

*header:* biblFull idno

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specGrp specGrpRef tag val

*textcrit:* listApp listWit

*textstructure:* floatingText

*transcr:* am ex subst

character data

*Note* May contain character data, phrase-level elements, and inter-level elements.

*Example*

```
<fDecl name="INV">
  <fDescr>inverted sentence</fDescr>
  <vRange>
    <vAlt>
      <binary value="true"/>
      <binary value="false"/>
    </vAlt>
  </vRange>
  <vDefault>
    <binary value="false"/>
  </vDefault>
</fDecl>
```

*Content model*

```
<content>
 <macroRef key="macro.limitedContent"/>
</content>
```

*Schema Declaration*

```
element fDescr
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   macro.limitedContent}
```

---

**<figDesc>** (description of figure) contains a brief prose description of the appearance or content of a graphic figure, for use when documenting an image without displaying it. [*15.4. Specific Elements for Graphic Images*]

*Module* figures — Tables, Formulæ, Graphics, and Notated Music

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by:* figure

*May contain*

*core:* abbr address bibl biblStruct choice cit date desc distinct email emph expan foreign gloss hi label list

listBibl measure measureGrp mentioned name num ptr q quote ref rs said soCalled stage term time title unit

*dictionaries:* lang

*drama:* camera caption castList move sound tech view

*figures:* table

*header:* biblFull idno

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specGrp specGrpRef tag val

*textcrit:* listApp listWit

*textstructure:* floatingText

*transcr:* am ex subst

character data

*Note* This element is intended for use as an alternative to the content of its parent <figure> element ; for example, to display when the image is required but the equipment in use cannot display graphic images. It may also be used for indexing or documentary purposes.

*Example*

```
<figure>
  <graphic  url="emblem1.png"/>
  <head>Emblemi d'Amore</head>
  <figDesc>A pair of naked winged cupids, each holding a
     flaming torch, in a rural setting.</figDesc>
</figure>
```

*Content model*

```
<content>
 <macroRef key="macro.limitedContent"/>
</content>
```

*Schema Declaration*

```
element figDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.limitedContent}
```

**<figure>** (figure) groups elements representing or containing graphic information such as an illustration, formula, or figure. [*15.4. Specific Elements for Graphic Images*]

*Module* figures — Tables, Formulæ, Graphics, and Notated Music

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.placement
  - *@place*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*

*Member of* model.global

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*gaiji:* char glyph

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct binaryObject cb cit desc ellipsis gap gb graphic head index l label lb lg list listBibl media meeting milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figDesc figure formula notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline docAuthor docDate epigraph floatingText postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Example*

```
<figure>
  <head>The View from the Bridge</head>
  <figDesc>A Whistleresque view showing four or five sailing boats in the
foreground, and a
    series of buoys strung out between them.</figDesc>
  <graphic  url="http://www.example.org/fig1.png"
   scale="0.5"/>
</figure>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.headLike"/>
  <classRef key="model.common"/>
  <elementRef key="figDesc"/>
  <classRef key="model.graphicLike"/>
  <classRef key="model.global"/>
  <classRef key="model.divBottom"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element figure
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.placement.attributes,
   att.typed.attributes,
   att.written.attributes,
   (
      model.headLike    | model.common    | figDesc    | model.graphicLike    | model.glob
}
```

**<fileDesc>** (file description) contains a full bibliographic description of an electronic file. [*2.2. The File Description 2.1.1. The TEI Header and Its Components*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by:* biblFull teiHeader

*May contain*

*header:* editionStmt extent notesStmt publicationStmt seriesStmt sourceDesc titleStmt

*Note* The major source of information for those seeking to create a catalogue entry or bibliographic citation for an electronic file. As such, it provides a title and statements of responsibility together with details of the publication or distribution of the file, of any series to which it belongs, and detailed bibliographic notes for matters not addressed elsewhere in the header. It also contains a full bibliographic description for the source or sources from which the electronic text was derived.

*Example*

```
<fileDesc>
  <titleStmt>
    <title>The shortest possible TEI document</title>
  </titleStmt>
  <publicationStmt>
    <p>Distributed as part of TEI P5</p>
  </publicationStmt>
  <sourceDesc>
    <p>No print source exists: this is an original digital text</p>
  </sourceDesc>
</fileDesc>
```

*Content model*

```
<content>
 <sequence>
  <sequence>
   <elementRef key="titleStmt"/>
   <elementRef key="editionStmt"
    minOccurs="0"/>
   <elementRef key="extent" minOccurs="0"/>
   <elementRef key="publicationStmt"/>
   <elementRef key="seriesStmt"
    minOccurs="0" maxOccurs="unbounded"/>
   <elementRef key="notesStmt"
    minOccurs="0"/>
  </sequence>
  <elementRef key="sourceDesc"
   minOccurs="1" maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element fileDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      (
         titleStmt,
         editionStmt?,
         extent?,
         publicationStmt,
         seriesStmt*,
         notesStmt?
      ),
      sourceDesc+
   )
}
```

**<filiation>** (filiation) contains information concerning the manuscript or other object's *filiation*, i.e. its relationship to other surviving manuscripts or other objects of the same text or contents, its *protographs*, *antigraphs* and *apographs*. [*11.6.1. The msItem and msItemStruct Elements*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.msItemPart

*Contained by*

*msdescription:* msItem msItemStruct

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<msContents>
  <msItem>
    <title>Beljakovski sbornik</title>
    <filiation type="protograph">Bulgarian</filiation>
    <filiation type="antigraph">Middle Bulgarian</filiation>
    <filiation type="apograph">
      <ref target="#DN17">Dujchev N 17</ref>
    </filiation>
  </msItem>
</msContents>
<!-- ... -->
<msDesc xml:id="DN17">
  <msIdentifier>
    <settlement>Faraway</settlement>
  </msIdentifier>
<!-- ... -->
</msDesc>
```

In this example, the reference to 'Dujchev N17' includes a link to some other manuscript description which has the identifier `DN17`.

*Example*

```
<msItem>
  <title>Guan-ben</title>
  <filiation>
    <p>The "Guan-ben" was widely current among mathematicians in the
        Qing dynasty, and "Zhao Qimei version" was also read. It is
        therefore difficult to know the correct filiation path to follow.
        The study of this era is much indebted to Li Di. We explain the
        outline of his conclusion here. Kong Guangsen
        (1752-1786)(17) was from the same town as Dai Zhen, so he obtained
        "Guan-ben" from him and studied it(18). Li Huang (d. 1811)
        (19) took part in editing Si Ku Quan Shu, so he must have had
        "Guan-ben". Then Zhang Dunren (1754-1834) obtained this version,
        and studied "Da Yan Zong Shu Shu" (The General Dayan
        Computation). He wrote Jiu Yi Suan Shu (Mathematics
        Searching for One, 1803) based on this version of Shu Xue Jiu
        Zhang (20).</p>
    <p>One of the most important persons in restoring our knowledge
        concerning the filiation of these books was Li Rui (1768(21)
        -1817)(see his biography). ... only two volumes remain of this
        manuscript, as far as chapter 6 (chapter 3 part 2) p.13, that is,
        question 2 of "Huan Tian San Ji" (square of three loops),
        which later has been lost.</p>
  </filiation>
</msItem>
<!--http://www2.nkfust.edu.tw/~jochi/ed1.htm-->
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element filiation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   macro.specialPara}
```

---

## <finalRubric>

<finalRubric> (final rubric) contains the string of words that denotes the end of a text division, often with an assertion as to its author and title, usually set off from the text itself by red ink, by a different size or type of script, or by some other such visual device. [*11.6.1. The msItem and msItemStruct Elements*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.msExcerpt
  - *@defective*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.msQuoteLike

*Contained by*

*msdescription:* msItem msItemStruct

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic
*gaiji:* g
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef specDesc specList tag val
*textcrit:* app witDetail
*textstructure:* floatingText
*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo
*verse:* caesura rhyme

character data

*Example*

```
<finalRubric>Explicit le romans de la Rose ou l'art
d'amours est toute enclose.</finalRubric>
<finalRubric>ok lúkv ver þar Brennu-Nials savgv</finalRubric>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element finalRubric
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.msExcerpt.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<fLib>** (feature library) assembles a library of <f> (feature) elements. [*19.4. Feature Libraries and Feature-Value Libraries*]

*Module* iso-fs — Feature Structures
*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*

*Member of* model.fsdDeclPart model.global.meta

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*iso-fs:* fsdDecl

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*iso-fs:* f

*Note* The global *n* attribute may be used to supply an informal name to categorize the library's contents.

*Example*

```
<fLib n="agreement features">
  <f xml:id="pers1" name="person">
    <symbol value="first"/>
  </f>
  <f xml:id="pers2" name="person">
    <symbol value="second"/>
  </f>
<!-- ... -->
  <f xml:id="nums" name="number">
    <symbol value="singular"/>
  </f>
  <f xml:id="nump" name="number">
    <symbol value="plural"/>
  </f>
<!-- ... -->
</fLib>
```

*Content model*

```
<content>
 <elementRef key="f" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element fLib
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   f+
}
```

---

## <floatingText>

<floatingText> (floating text) contains a single text of any kind, whether unitary or composite, which interrupts the text containing it at any point and after which the surrounding text resumes. [*4.3.2. Floating Texts*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*

- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.declaring
  - *@decls*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.attributable

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author biblScope cit citedRange corr del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled sp speaker stage street term textLang title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* actor camera caption castList epilogue performance prologue role roleDesc set sound tech view

*figures:* cell figDesc figure

*header:* change distributor edition extent geoDecl handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* cb ellipsis gap gb index lb milestone note noteGrp pb

*figures:* figure notatedMusic

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*textstructure:* back body front group

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* A floating text has the same content as any other <text> and may thus be interrupted by another floating text, or contain a <group> of tesselated texts.

*Example*

```
<body>
  <div  type="scene">
    <sp>
      <p>Hush, the players begin...</p>
    </sp>
    <floatingText  type="pwp">
      <body>
        <div  type="act">
          <sp>
            <l>In Athens our tale takes place [...]</l>
          </sp>
<!-- ... rest of nested act here -->
        </div>
      </body>
    </floatingText>
    <sp>
      <p>Now that the play is finished ...</p>
    </sp>
  </div>
</body>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.global"
   minOccurs="0" maxOccurs="unbounded"/>
  <sequence minOccurs="0">
   <elementRef key="front"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <alternate>
   <elementRef key="body"/>
   <elementRef key="group"/>
  </alternate>
  <classRef key="model.global"
   minOccurs="0" maxOccurs="unbounded"/>
  <sequence minOccurs="0">
   <elementRef key="back"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element floatingText
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declaring.attributes,
   att.typed.attributes,
   (
      model.global*,
      ( ( front, model.global* )? ),
      ( body | group ),
      model.global*,
      ( ( back, model.global* )? )
   )
}
```

---

**<floruit>** (floruit) contains information about a person's period of activity. [*14.3.2.1. Personal Characteristics*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*

    * *@max*
    * *@confidence*
  - att.editLike
    - *@evidence*
    - *@instant*

*Member of* model.persStateLike

*Contained by*

*namesdates:* person personGrp persona

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<floruit notBefore="1066" notAfter="1100"/>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element floruit
{
   att.global.attributes,
   att.global.analytic.attributes,
```

```
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   macro.phraseSeq}
```

---

**<foliation>** (foliation) describes the numbering system or systems used to count the leaves or pages in a codex or similar object. [*11.7.1.4. Foliation*]

*Module* msdescription — Manuscript Description

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility

* *@cert*
* *@resp*

– att.global.source
* *@source*

*Contained by* *msdescription:* supportDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<foliation>Contemporary foliation in red
roman numerals in the centre
of the outer margin.</foliation>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element foliation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.specialPara}
```

---

**<foreign>** (foreign) identifies a word or phrase as belonging to some language other than that of the surrounding text. [*3.3.2.1. Foreign Words or Expressions*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*

  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*

*Member of* model.emphLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The global *xml:lang* attribute should be supplied for this element to identify the language of the word or phrase marked. As elsewhere, its value should be a language tag as defined in *vi.1 Language Identification*.

This element is intended for use only where no other element is available to mark the phrase or words concerned. The global *xml:lang* attribute should be used in preference to this element where it is intended to mark the language of the whole of some text element.

The <distinct> element may be used to identify phrases belonging to sublanguages or registers not generally regarded as true languages.

*Example*

```
This is
heathen Greek to you still? Your <foreign  xml:lang="la">lapis
philosophicus</foreign>?
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element foreign
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   macro.phraseSeq}
```

---

## <forename>

(forename) contains a forename, given or baptismal name. [*14.2.1. Personal Names*]

*Module* namesdates — Names, Dates, People, and Places
*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.personal
  - *@full*
  - *@sort*

- att.naming
  - *@role*
  - *@nymRef*

  - att.canonical
    - *@key*
    - *@ref*

- att.typed
  - *@type*
  - *@subtype*

*Member of* model.persNamePart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<persName>
  <roleName>Ex-President</roleName>
  <forename>George</forename>
  <surname>Bush</surname>
</persName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element forename
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.personal.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

## <forest> (forest) provides for groups of rooted trees. [*20.3. Another Tree Notation*]

*Module* nets — Graphs, Networks, and Trees

*Attributes*

- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.typed
  - *@type*
  - *@subtype*

*Member of* model.divPart model.standOffPart

*Contained by*

*core:* item note q quote said stage textLang

*drama:* castList epilogue performance prologue set view

*figures:* cell figure

*header:* change handNote licence scriptNote

*linking:* standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*nets:* listForest

*tagdocs:* specGrp

*textcrit:* lem rdg

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript
*transcr:* metamark
*May contain*
*nets:* eTree tree triangle
*Note* One or more trees, embedding trees, or underspecified embedding trees (triangles).
*Example*

```
<forest n="ex5" type="derivation-syntactic">
  <eTree n="Stage 1" xml:id="s1SBAR">
    <label>S'</label>
    <eTree xml:id="s1S">
      <label>S</label>
      <eTree xml:id="s1NP1">
        <label>NP</label>
        <eLeaf>
          <label>you</label>
        </eLeaf>
      </eTree>
      <eTree xml:id="s1VP">
        <label>VP</label>
        <eTree xml:id="s1V">
          <label>V</label>
          <eLeaf>
            <label>do</label>
          </eLeaf>
        </eTree>
      </eTree>
    </eTree>
  </eTree>
  <eTree n="Stage 2" xml:id="s2SBAR"
    corresp="#s1SBAR">
    <label>S'</label>
    <eTree xml:id="s2S" corresp="#s1S">
      <label>S</label>
      <eTree xml:id="s2NP1" copyOf="#s1NP1">
        <label>NP</label>
      </eTree>
      <eTree xml:id="s2VP" corresp="#s1VP">
        <label>VP</label>
        <eTree xml:id="s2V" copyOf="#s1V">
          <label>V</label>
        </eTree>
        <eTree xml:id="s2NP2" corresp="#s1NP2">
          <label>NP</label>
          <eLeaf corresp="#s1WH">
            <label>t</label>
          </eLeaf>
        </eTree>
      </eTree>
    </eTree>
  </eTree>
</forest>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
```

```
  <elementRef key="tree"/>
  <elementRef key="eTree"/>
  <elementRef key="triangle"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element forest
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   ( tree | eTree | triangle )+
}
```

---

**<form>** (form information group) groups all the information on the written and spoken forms of one headword. [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

• att.lexicographic
– *@expand*
– *@split*
– *@value*
– *@location*
– *@mergedIn*
– *@opt*

– att.datcat
* *@datcat*
* *@valueDatcat*
* *@targetDatcat*

– att.lexicographic.normalized
* *@norm*
* *@orig*

• att.typed
– ~~type~~
– @subtype

`@type` classifies form as simple, compound, etc.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **simple** single free lexical item
**lemma** the headword itself
**variant** a variant form
**compound** word formed from simple lexical items
**derivative** word derived from headword
**inflected** word in other than usual dictionary form
**phrase** multiple-word lexical item

*Member of* model.entryPart model.entryPart.top model.formPart
*Contained by*
*core:* cit
*dictionaries:* dictScrap entry entryFree form hom re sense superEntry
*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index label lb list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* case colloc form gen gram gramGrp hyph iType lang lbl mood number oRef orth pRef per pos pron stress subc syll tns usg

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<form>
  <orth>zaptié</orth>
  <orth>zaptyé</orth>
</form>
```

(from TLFi)

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.formPart"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element form
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.typed.attribute.subtype,
   attribute type
   {
      "simple"
    | "lemma"
    | "variant"
    | "compound"
    | "derivative"
    | "inflected"
    | "phrase"
   }?,
   (
      text
    | model.gLike    | model.phrase    | model.inter    | model.formPart    | model.global
}
```

---

**<formula>** (formula) contains a mathematical or other formula. [*15.2. Formulæ and Mathematical Expressions*]

*Module* figures — Tables, Formulæ, Graphics, and Notated Music

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*

    - * *@copyOf*
    - * *@next*
    - * *@prev*
    - * *@exclude*
    - * *@select*
    - – att.global.rendition
    - * *@rend*
    - * *@style*
    - * *@rendition*
    - – att.global.responsibility
    - * *@cert*
    - * *@resp*
    - – att.global.source
    - * *@source*
  - • att.cmc
    - – *@generatedBy*
  - • att.notated
    - – *@notation*

*Member of* model.graphicLike

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figure formula table

*gaiji:* char glyph

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage facsimile fw metamark mod restore retrace secl sourceDoc supplied surface surplus zone

*verse:* rhyme

*May contain*

*core:* binaryObject graphic hi media q

*figures:* formula

character data

*Example*

```
<formula notation="tex">$E=mc^2$</formula>
```

*Example*

```
<formula notation="none">E=mc<hi rend="sup">2</hi>
</formula>
```

*Example*

```
<formula notation="mathml">
  <m:math>
    <m:mi>E</m:mi>
    <m:mo>=</m:mo>
    <m:mi>m</m:mi>
    <m:msup>
      <m:mrow>
        <m:mi>c</m:mi>
      </m:mrow>
      <m:mrow>
        <m:mn>2</m:mn>
      </m:mrow>
    </m:msup>
  </m:math>
</formula>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.graphicLike"/>
  <classRef key="model.hiLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element formula
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.notated.attributes,
```

```
   ( text | model.graphicLike | model.hiLike )*
}
```

---

**<front>** (front matter) contains any prefatory matter (headers, abstracts, title page, prefaces, dedications, etc.) found at the start of a document, before the main body. [*4.6. Title Pages 4. Default Text Structure*]

*Module* textstructure — Default Text Structure

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declaring
  - *@decls*

*Contained by:* textstructure: floatingText text
*transcr:* facsimile

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* cb divGen ellipsis gap gb head index lb listBibl meeting milestone note noteGrp p pb

*drama:* castList epilogue performance prologue set

*figures:* figure notatedMusic

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* schemaSpec

*textcrit:* app witDetail

*textstructure:* argument byline closer dateline div div1 docAuthor docDate docEdition docImprint docTitle epigraph postscript salute signed titlePage titlePart trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* Because cultural conventions differ as to which elements are grouped as front matter and which as back matter, the content models for the <front> and <back> elements are identical.

*Example*

```
<front>
  <epigraph>
    <quote>Nam Sibyllam quidem Cumis ego ipse oculis meis vidi in ampulla
        pendere, et cum illi pueri dicerent: <q  xml:lang="grc">Σίβυλλα τί
           θέλεις</q>; respondebat illa: <q  xml:lang="grc">ἀποθανεῖν θέλω.</q>
    </quote>
  </epigraph>
  <div  type="dedication">
    <p>For Ezra Pound <q  xml:lang="it">il miglior fabbro.</q>
    </p>
  </div>
</front>
```

*Example*

```
<front>
  <div  type="dedication">
    <p>To our three selves</p>
  </div>
  <div  type="preface">
    <head>Author's Note</head>
    <p>All the characters in this book are purely imaginary, and if the
        author has used names that may suggest a reference to living persons
        she has done so inadvertently. ...</p>
  </div>
</front>
```

*Example*

```
<front>
  <div  type="abstract">
    <div>
      <head> BACKGROUND:</head>
      <p>Food insecurity can put children at greater risk of obesity because
            of altered food choices and nonuniform consumption patterns.</p>
    </div>
```

```
      <div>
        <head> OBJECTIVE:</head>
        <p>We examined the association between obesity and both child-level
              food insecurity and personal food insecurity in US children.</p>
      </div>
      <div>
        <head> DESIGN:</head>
        <p>Data from 9,701 participants in the National Health and Nutrition
              Examination Survey, 2001-2010, aged 2 to 11 years were analyzed.
              Child-level food insecurity was assessed with the US Department of
              Agriculture's Food Security Survey Module based on eight
              child-specific questions. Personal food insecurity was assessed with
              five additional questions. Obesity was defined, using physical
              measurements, as body mass index (calculated as kg/m2) greater than
              or equal to the age- and sex-specific 95th percentile of the Centers
              for Disease Control and Prevention growth charts. Logistic
              regressions adjusted for sex, race/ethnic group, poverty level, and
              survey year were conducted to describe associations between obesity
              and food insecurity.</p>
      </div>
      <div>
        <head> RESULTS:</head>
        <p>Obesity was significantly associated with personal food insecurity
              for children aged 6 to 11 years (odds ratio=1.81; 95% CI 1.33 to
              2.48), but not in children aged 2 to 5 years (odds ratio=0.88; 95%
              CI 0.51 to 1.51). Child-level food insecurity was not associated
              with obesity among 2- to 5-year-olds or 6- to 11-year-olds.</p>
      </div>
      <div>
        <head> CONCLUSIONS:</head>
        <p>Personal food insecurity is associated with an increased risk of
              obesity only in children aged 6 to 11 years. Personal
              food-insecurity measures may give different results than aggregate
              food-insecurity measures in children.</p>
      </div>
    </div>
  </front>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.frontPart"/>
   <classRef key="model.pLike"/>
   <classRef key="model.pLike.front"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="0">
   <alternate>
    <sequence>
     <classRef key="model.div1Like"/>
     <alternate minOccurs="0"
      maxOccurs="unbounded">
      <classRef key="model.div1Like"/>
      <classRef key="model.frontPart"/>
      <classRef key="model.global"/>
```

```
     </alternate>
    </sequence>
    <sequence>
     <classRef key="model.divLike"/>
     <alternate minOccurs="0"
      maxOccurs="unbounded">
      <classRef key="model.divLike"/>
      <classRef key="model.frontPart"/>
      <classRef key="model.global"/>
     </alternate>
    </sequence>
   </alternate>
   <sequence minOccurs="0">
    <classRef key="model.divBottom"/>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <classRef key="model.divBottom"/>
     <classRef key="model.global"/>
    </alternate>
   </sequence>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element front
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   (
      ( model.frontPart | model.pLike | model.pLike.front | model.global )*,
      (
         (
            (
               (
                  model.div1Like,
                  ( model.div1Like | model.frontPart | model.global )*
               )
             | (
                  model.divLike,
                  ( model.divLike | model.frontPart | model.global )*
               )
            ),
            ( ( model.divBottom, ( model.divBottom | model.global )* )? )
         )?
      )
   )
}
```

**<fs>** (feature structure) represents a *feature structure*, that is, a collection of feature-value pairs organized as a structural unit. [*19.2. Elementary Feature Structures and the Binary Feature Value*]

*Module* iso-fs — Feature Structures

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*

`@type` specifies the type of the feature structure.
*Status* Optional

*Datatype* teidata.enumerated

`@feats` (features) references the feature-value specifications making up this feature structure.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Note* May be used either instead of having features as content, or in addition. In the latter case, the features referenced and contained are unified.

*Member of* model.featureVal.complex model.global.meta

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*iso-fs:* bicond cond f fvLib if vAlt vColl vDefault vLabel vMerge vNot vRange

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*iso-fs:* f

*Example*

```
<fs type="agreement_structure">
  <f name="person">
    <symbol value="third"/>
```

```
    </f>
    <f  name="number">
      <symbol  value="singular"/>
    </f>
  </fs>
```

*Content model*

```
<content>
 <elementRef key="f" minOccurs="0"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element fs
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datcat.attributes,
   attribute type { text }?,
   attribute feats { list { + } }?,
   f*
}
```

---

## <fsConstraints>

(feature-structure constraints) specifies constraints on the content of valid feature structures. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Contained by:* *iso-fs:* fsDecl

*May contain*

*iso-fs:* bicond cond

*Note* May contain a series of conditional or biconditional elements.

*Example*

```
<fsConstraints>
  <cond>
    <fs>
<!-- ...-->
    </fs>
    <then/>
    <fs>
<!-- ... -->
    </fs>
  </cond>
</fsConstraints>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <elementRef key="cond"/>
  <elementRef key="bicond"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element fsConstraints
{
```

```
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( cond | bicond )*
}
```

---

**<fsdDecl>** (feature system declaration) provides a feature system declaration comprising one or more feature structure declarations or feature structure declaration links. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

* *@source*

*Member of* model.encodingDescPart model.resource
*Contained by*
*core:* teiCorpus
*header:* encodingDesc
*textstructure:* TEI
*May contain*
*iso-fs:* fLib fsDecl fsdLink fvLib
*Example*

```
<fsdDecl>
  <fsDecl  type="GPSG">
    <fDecl  xml:id="GPSG-1"  name="GPSG_feat1">
      <vRange>
        <vAlt>
          <symbol  value="red"/>
          <symbol  value="blue"/>
          <symbol  value="green"/>
        </vAlt>
      </vRange>
    </fDecl>
<!--other feature declarations for GPSG here ... -->
  </fsDecl>
  <fsdLink  type="subentry"
    target="http://www.example.com/fsdLib.xml#LX123"/>
</fsdDecl>
```

*Content model*

```
<content>
 <classRef key="model.fsdDeclPart"
  minOccurs="1" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element fsdDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   model.fsdDeclPart+
}
```

---

**<fsDecl>** (feature structure declaration) declares one type of feature structure. [*19.11. Feature System Declaration*]
*Module* iso-fs — Feature Structures
*Attributes* • att.global

  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*

`@type` gives a name for the type of feature structure being declared.

*Status* Required

*Datatype* teidata.enumerated

`@baseTypes` gives the name of one or more typed feature structures from which this type inherits feature specifications and constraints; if this type includes a feature specification with the same name as that of any of those specified by this attribute, or if more than one specification of the same name is inherited, then the set of possible values is defined by unification. Similarly, the set of constraints applicable is derived by combining those specified explicitly within this element with those implied by the *baseTypes* attribute. When no *baseTypes* attribute is specified, no feature specification or constraint is inherited.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.name separated by whitespace

*Note* Inheritance is defined here as a monotonous relation.

The process of combining constraints may result in a contradiction, for example if two specifications for the same feature specify disjoint ranges of values, and at least one such specification is mandatory. In such a case, there is no valid representative for the type being defined.

*Member of* model.fsdDeclPart

*Contained by*

*iso-fs:* fsdDecl

*May contain*

*iso-fs:* fDecl fsConstraints fsDescr

*Example*

```
<fsDecl  type="SomeName">
  <fsDescr>Describes what this type of fs represents</fsDescr>
  <fDecl  name="featureOne">
<!-- The declaration for featureOne -->
    <vRange>
<!-- the range of possible values for this feature -->
    </vRange>
  </fDecl>
  <fDecl  name="featureTwo">
<!-- The declaration for featureTwo -->
    <vRange>
<!-- the range of possible values for this feature -->
    </vRange>
  </fDecl>
  <fsConstraints>
<!-- Any additional constraints for the feature structure -->
  </fsConstraints>
</fsDecl>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="fsDescr" minOccurs="0"/>
  <elementRef key="fDecl" minOccurs="1"
   maxOccurs="unbounded"/>
  <elementRef key="fsConstraints"
   minOccurs="0"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element fsDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
```

```
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datcat.attributes,
   attribute type { text },
   attribute baseTypes { list { + } }?,
   ( fsDescr?, fDecl+, fsConstraints? )
}
```

---

**<fsDescr>** (feature system description (in FSD)) describes in prose what is represented by the type of feature structure declared in the enclosing fsDecl. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by:* *isd-lfs:* fsDecl

*May contain*

*core:* abbr address bibl biblStruct choice cit date desc distinct email emph expan foreign gloss hi label list listBibl measure measureGrp mentioned name num ptr q quote ref rs said soCalled stage term time title unit

*dictionaries:* lang

*drama:* camera caption castList move sound tech view

*figures:* table

*header:* biblFull idno

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specGrp specGrpRef tag val

*textcrit:* listApp listWit

*textstructure:* floatingText

*transcr:* am ex subst

character data

*Note* May contain character data, phrase-level elements, and inter-level elements.

*Example*

```
<fsDecl type="Agreement">
  <fsDescr>This type of feature structure encodes the features
     for subject-verb agreement in English</fsDescr>
  <fDecl name="PERS">
    <fDescr>person (first, second, or third)</fDescr>
    <vRange>
      <vAlt>
        <symbol value="first"/>
        <symbol value="second"/>
        <symbol value="third"/>
      </vAlt>
    </vRange>
  </fDecl>
  <fDecl name="NUM">
    <fDescr>number (singular or plural)</fDescr>
    <vRange>
      <vAlt>
        <symbol value="singular"/>
        <symbol value="plural"/>
      </vAlt>
    </vRange>
  </fDecl>
</fsDecl>
```

*Content model*

```
<content>
 <macroRef key="macro.limitedContent"/>
</content>
```

*Schema Declaration*

```
element fsDescr
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.limitedContent}
```

---

**<fsdLink>** (feature structure declaration link) associates the name of a typed feature structure with a feature structure declaration for it. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*

    * *@resp*
  - att.global.source
    * *@source*

`@type` identifies the type of feature structure to be documented; this will be the value of the *type* attribute on at least one feature structure.
*Status* Required
*Datatype* teidata.enumerated

`@target` supplies a pointer to a feature structure declaration (<fsDecl>) element within the current document or elsewhere.
*Status* Required
*Datatype* teidata.pointer

*Member of* model.fsdDeclPart
*Contained by*
*iso-fs:* fsdDecl
*May contain* Empty element
*Example*

```
<fsdLink type="subentry"
 target="http://www.example.com/fsdLib.xml#L1234"/>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element fsdLink
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute type { text },
   attribute target { text },
   empty
}
```

---

## <funder>

(funding body) specifies the name of an individual, institution, or organization responsible for the funding of a project or text. [*2.2.1. The Title Statement*]

*Module* header — The TEI Header
*Attributes* • att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic

    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*

- att.datable.w3c
  - *@when*
  - *@notBefore*
  - *@notAfter*
  - *@from*
  - *@to*

*Member of* model.respLike

*Contained by*

*core:* bibl monogr

*header:* editionStmt titleStmt

*msdescription:* msItem

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin

character data

*Note* Funders provide financial support for a project; they are distinct from *sponsors* (see element <sponsor>), who provide intellectual support and authority.

*Example*

```
<funder>The National Endowment for the Humanities, an independent federal
agency</funder>
<funder>Directorate General XIII of the Commission of the European
Communities</funder>
<funder>The Andrew W. Mellon Foundation</funder>
<funder>The Social Sciences and Humanities Research Council of Canada</funder>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element funder
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   macro.phraseSeq.limited}
```

---

**<fvLib>** (feature-value library) assembles a library of reusable feature value elements (including complete feature structures). [*19.4. Feature Libraries and Feature-Value Libraries*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*

    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*

*Member of* model.fsdDeclPart model.global.meta

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*iso-fs:* fsdDecl

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*iso-fs:* binary default fs numeric string symbol vAlt vColl vLabel vMerge vNot

*Note* A feature value library may include any number of values of any kind, including multiple occurrences of

identical values such as <binary value="true"/> or default. The only thing guaranteed unique in a feature value library is the set of labels used to identify the values.

*Example*

```
<fvLib n="symbolic values">
  <symbol xml:id="sfirst" value="first"/>
  <symbol xml:id="ssecond" value="second"/>
<!-- ... -->
  <symbol xml:id="ssing" value="singular"/>
  <symbol xml:id="splur" value="plural"/>
<!-- ... -->
</fvLib>
```

*Content model*

```
<content>
 <classRef key="model.featureVal"
  minOccurs="0" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element fvLib
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   model.featureVal*
}
```

---

**<fw>** (forme work) contains a running head (e.g. a header, footer), catchword, or similar material appearing on the current page. [*12.6. Headers, Footers, and Similar Matter*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.placement
  - *@place*
- att.written
  - *@hand*
- att.typed
  - ~~type~~
  - @subtype

`@type` classifies the material encoded according to some useful typology.

*Derived from* att.typed

*Status* Recommended

*Datatype* teidata.enumerated

*Sample values include:* **header** a running title at the top of the page

**footer** a running title at the bottom of the page

**pageNum** (page number) a page number or foliation symbol

**lineNum** (line number) a line number, either of prose or poetry

**sig** (signature) a signature or gathering symbol

**catch** (catchword) a catch-word

*Member of* model.milestoneLike

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list listBibl measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc subst supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Where running heads are consistent throughout a chapter or section, it is usually more convenient to relate them to the chapter or section, e.g. by use of the *rend* attribute. The <fw> element is intended for cases where the running head changes from page to page, or where details of page layout and the internal structure of the running head are of paramount importance.

*Example*

```
<fw type="sig" place="bottom">C3</fw>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element fw
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.placement.attributes,
   att.written.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.phraseSeq}
```

---

**<g>** (character or glyph) represents a glyph, or a non-standard character. [*5. Characters, Glyphs, and Writing Modes*]

*Module* gaiji — Characters, Glyphs, and Writing Modes

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs

    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.typed
  - *@type*
  - *@subtype*

`@ref` points to a description of the character or glyph intended.
*Status* Optional
*Datatype* teidata.pointer

*Member of* model.gLike

*Contained by*

*analysis:* c cl interp m pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure measureGrp mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said series sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number oRef orth pRef per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*gaiji:* mapping

*header:* change distributor edition extent geoDecl handNote idno licence scriptNote

*iso-fs:* string

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote depth dim explicit filiation finalRubric foliation height heraldry incipit layout locus material msName musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark width

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg memberOf

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage ex fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain* Character data only

*Note* The name `g` is short for *gaiji*, which is the Japanese term for a non-standardized character or glyph.

*Example*

```
<g ref="#ctlig">ct</g>
```

This example points to a <glyph> element with the identifier `ctlig` like the following:

```
<glyph xml:id="ctlig">
<!-- here we describe the particular ct-ligature intended -->
</glyph>
```

*Example*

```
<g ref="#per-glyph">per</g>
```

The medieval brevigraph per could similarly be considered as an individual glyph, defined in a <glyph> element with the identifier `per-glyph` as follows:

```
<glyph xml:id="per-glyph">
<!-- ... -->
</glyph>
```

*Content model* `<content> <textNode/></content>`

*Schema Declaration*

```
element g
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   attribute ref { text }?,
   text
}
```

---

**<gap>** (gap) indicates a point where material has been omitted in a transcription, whether for editorial reasons

described in the TEI header, as part of sampling practice, or because the material is illegible, invisible, or inaudible. [*3.5.3. Additions, Deletions, and Omissions*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*
  - att.ranging

* *@atLeast*
* *@atMost*
* *@min*
* *@max*
* *@confidence*

- att.editLike
  - *@evidence*
  - *@instant*
- att.timed
  - *@start*
  - *@end*

  - att.duration
  * att.duration.iso
  · *@dur-iso*
  * att.duration.w3c
  · *@dur*

`@reason` (reason) gives the reason for omission.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace
*Suggested values include:* **cancelled** (cancelled)
**deleted** (deleted)
**editorial** (editorial) for features omitted from transcription due to editorial policy
**illegible** (illegible)
**inaudible** (inaudible)
**irrelevant** (irrelevant)
**sampling** (sampling)

`@agent` (agent) in the case of text omitted because of damage, categorizes the cause of the damage, if it can be identified.
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **rubbing** (rubbing) damage results from rubbing of the leaf edges
**mildew** (mildew) damage results from mildew on the leaf surface
**smoke** (smoke) damage results from smoke

*Member of* model.global.edit
*Contained by*
*analysis:* cl m phr s span w
*cmc:* post
*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit
*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose
*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*certainty:* certainty precision respons

*core:* desc

*Note* The <gap>, <unclear>, and <del> core tag elements may be closely allied in use with the <damage> and <supplied> elements, available when using the additional tagset for transcription of primary sources. See section *12.3.3.2. Use of the gap, del, damage, unclear, and supplied Elements in Combination* for discussion of which element is appropriate for which circumstance.

The <gap> tag simply signals the editors decision to omit or inability to transcribe a span of text. Other information, such as the interpretation that text was deliberately erased or covered, should be indicated using the relevant tags, such as <del> in the case of deliberate deletion.

*Example*

```
<gap quantity="4" unit="chars"
  reason="illegible"/>
```

*Example*

```
<gap quantity="1" unit="essay"
  reason="sampling"/>
```

*Example*

```
<del>
  <gap atLeast="4" atMost="8" unit="chars"
    reason="illegible"/>
</del>
```

*Example*

```
<gap extent="several lines" reason="lost"/>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.descLike"/>
  <classRef key="model.certLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element gap
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   attribute reason
   {
      list
      {
         (
            "cancelled"
          | "deleted"
          | "editorial"
          | "illegible"
          | "inaudible"
          | "irrelevant"
          | "sampling"
         )+
      }
   }?,
   attribute agent { text }?,
   ( model.descLike | model.certLike )*
}
```

---

**<gb>** (gathering beginning) marks the beginning of a new gathering or quire in a transcribed codex. [*3.11.3. Milestone Elements*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.breaking
  - *@break*
- att.cmc
  - *@generatedBy*
- att.edition
  - *@ed*
  - *@edRef*
- att.spanning
  - *@spanTo*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.milestoneLike

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list listBibl measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc subst supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Note* By convention, <gb> elements should appear at the start of the first page in the gathering. The global *n* attribute indicates the number or other value used to identify this gathering in a collation.

The *type* attribute may be used to further characterize the gathering in any respect.

*Example*

```
<gb  n="iii"/>
<pb  n="2r"/>
<!-- material from page 2 recto of gathering iii here -->
<pb  n="2v"/>
<!-- material from page 2 verso of gathering iii here -->
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element gb
```

```
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.breaking.attributes,
   att.cmc.attributes,
   att.edition.attributes,
   att.spanning.attributes,
   att.typed.attributes,
   empty
}
```

---

**<gen>** (gender) identifies the morphological gender of a lexical item, as given in the dictionary. [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*

- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*
- att.lexicographic.normalized
  - *@norm*
  - *@orig*

*Member of* model.morphLike

*Contained by*

*dictionaries:* dictScrap entryFree form gramGrp

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May contain character data and phrase-level elements. Typical content will be *masculine*, *feminine*, *neuter* etc.

This element is synonymous with <gram type="gender">.

*Example*

```
<entry>
  <form>
    <orth>pamplemousse</orth>
  </form>
  <gramGrp>
    <pos>noun</pos>
    <gen>masculine</gen>
  </gramGrp>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element gen
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   macro.paraContent}
```

---

## <gender>

(gender) specifies the gender identity of a person, persona, or character. [*14.3.2.1. Personal Characteristics*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.datable
  - *@period*

- att.datable.custom
  - *@when-custom*
  - *@notBefore-custom*
  - *@notAfter-custom*
  - *@from-custom*
  - *@to-custom*
  - *@datingPoint*
  - *@datingMethod*
- att.datable.iso
  - *@when-iso*
  - *@notBefore-iso*
  - *@notAfter-iso*
  - *@from-iso*

  - * *@to-iso*
  - att.datable.w3c
  - * *@when*
  - * *@notBefore*
  - * *@notAfter*
  - * *@from*
  - * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.typed
  - *@type*
  - *@subtype*

`@value` supplies a coded value for gender identity.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.gender separated by whitespace

*Note* Values for this attribute may be locally defined by a project, or they may refer to an external standard.

*Member of* model.persStateLike

*Contained by*

*namesdates:* person personGrp persona

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* As with other culturally-constructed traits such as age and sex, the way in which this concept is described in different cultural contexts varies. The normalizing attributes are provided only as an optional means of simplifying that variety for purposes of interoperability or project-internal taxonomies for consistency, and should not be used where that is inappropriate or unhelpful. The content of the element may be used to describe the intended concept in more detail.

*Example*

```
<gender value="W">woman</gender>
```

*Example*

```
<gender value="NB">non-binary</gender>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element gender
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.typed.attributes,
   attribute value { list { + } }?,
   macro.phraseSeq}
```

---

**<genName>** (generational name component) contains a name component used to distinguish otherwise similar names on the basis of the relative ages or generations of the persons named. [*14.2.1. Personal Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.personal
  - *@full*
  - *@sort*

  - att.naming
    * *@role*
    * *@nymRef*

    * att.canonical
    · *@key*
    · *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.persNamePart
*Contained by*
*analysis:* cl phr s span
*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<persName>
  <forename>Charles</forename>
  <genName>II</genName>
</persName>
```

*Example*

```
<persName>
  <surname>Pitt</surname>
  <genName>the Younger</genName>
</persName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element genName
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.personal.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<geo>** (geographical coordinates) contains any expression of a set of geographic coordinates, representing a point, line, or area on the surface of the earth in some notation. [*14.3.4.1. Varieties of Location*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.declaring
  - *@decls*

*Member of* model.measureLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure measureGrp meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain* Character data only

*Note* Uses of <geo> can be associated with a coordinate system, defined by a <geoDecl> element supplied in the TEI header, using the *decls* attribute. If no such link is made, the assumption is that the content of each <geo> element will be a pair of numbers separated by whitespace, to be interpreted as latitude followed by longitude according to the World Geodetic System.

*Example*

```
<geoDecl  xml:id="WGS"  datum="WGS84">World Geodetic System</geoDecl>
<geoDecl  xml:id="OS"  datum="OSGB36">Ordnance Survey</geoDecl>
<!-- ... -->
<location>
  <desc>A tombstone plus six lines of
     Anglo-Saxon text, built into the west tower (on the south side
     of the archway, at 8 ft. above the ground) of the
     Church of St. Mary-le-Wigford in Lincoln.</desc>
  <geo  decls="#WGS">53.226658 -0.541254</geo>
  <geo  decls="#OS">SK 97481 70947</geo>
</location>
```

*Example*

```
<geo>41.687142 -74.870109</geo>
```

*Content model* `<content> <textNode/></content>`

*Schema Declaration*

```
element geo
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declaring.attributes,
```

```
   text
}
```

---

**<geoDecl>** (geographic coordinates declaration) documents the notation and the datum used for geographic coordinates expressed as content of the <geo> element elsewhere within the document. [*2.3.8. The Geographic Coordinates Declaration*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable
  - *@default*

`@datum` supplies a commonly used code name for the datum employed.

*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **WGS84** (World Geodetic System) a pair of numbers to be interpreted as latitude followed by longitude according to the World Geodetic System.[Default]

**MGRS** (Military Grid Reference System) values supplied follow the Military Grid Reference System, which designates grid zones in a string of letters and numbers that distinctly indicate each square meter on the planet.

**OSGB36** (ordnance survey great britain) the value supplied is to be interpreted as a British National Grid Reference.

**ED50** (European Datum coordinate system) the value supplied is to be interpreted as latitude followed by longitude according to the European Datum coordinate system.

*Member of* model.encodingDescPart

*Contained by*

*header:* encodingDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<geoDecl datum="OSGB36"/>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:geoDecl"/> </sch:pattern>

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
```

```
</content>
```

*Schema Declaration*

```
element geoDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   attribute datum { "WGS84" | "MGRS" | "OSGB36" | "ED50" }?,
   macro.phraseSeq}
```

---

## <geogFeat> (geographical feature name) contains a common noun identifying some geographical feature contained within a geographic name, such as valley, mount, etc. [*14.2.3. Place Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*

    * *@max*
    * *@confidence*
  - att.editLike
    - *@evidence*
    - *@instant*
  - att.naming
    - *@role*
    - *@nymRef*
  - att.canonical
    * *@key*
    * *@ref*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.offsetLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<geogName>The <geogFeat>vale</geogFeat> of White Horse</geogName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element geogFeat
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
```

```
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<geogName>** (geographical name) identifies a name associated with some geographical feature such as Windrush Valley or Mount Sinai. [*14.2.3. Place Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc

- *@generatedBy*

- att.datable
  - *@period*

  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*
    - *@to-custom*
    - *@datingPoint*
    - *@datingMethod*
  - att.datable.iso
    - *@when-iso*
    - *@notBefore-iso*
    - *@notAfter-iso*
    - *@from-iso*
    - *@to-iso*
  - att.datable.w3c
    - *@when*
    - *@notBefore*
    - *@notAfter*
    - *@from*
    - *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*

  - att.canonical
    - *@key*
    - *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.placeNamePart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num

orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage unitDef

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions altIdentifier catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msIdentifier musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectIdentifier objectName occupation offset org orgName persName persPronouns place placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<geogName>
  <geogFeat>Mount</geogFeat>
  <name>Sinai</name>
</geogName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element geogName
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<gi>** (element name) contains the name (generic identifier) of an element. [*23. Documentation Elements 23.5. Element Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

* *@ana*

– att.global.change
* *@change*

– att.global.facs
* *@facs*

– att.global.linking
* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@scheme` supplies the name of the scheme in which this name is defined.

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **TEI** this element is part of the TEI scheme.[Default]

**DBK** (docbook) this element is part of the Docbook scheme.

**XX** (unknown) this element is part of an unknown scheme.

**Schematron** this element is from Schematron.

**HTML** this element is from the HTML scheme.

*Member of* model.phrase.xml

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain* XSD Name

*Example*

```
<p>The <gi>xhtml:li</gi> element is roughly analogous to the <gi>item</gi>
element, as is the
<gi  scheme="DBK">listItem</gi> element.</p>
```

This example shows the use of both a namespace prefix and the *scheme* attribute as alternative ways of indicating that the <gi> in question is not a TEI element name: in practice only one method should be adopted.

*Content model* `<content> <dataRef key="teidata.name"/></content>`

*Schema Declaration*

```
element gi
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute scheme { text }?,
   teidata.name}
```

---

## <gloss>

(gloss) identifies a phrase or word used to provide a gloss or definition for some other word or phrase. [*3.4.1. Terms and Glosses 23.4.1. Description of Components*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cReferencing
  - *@cRef*
- att.cmc
  - *@generatedBy*
- att.declaring
  - *@decls*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*
- att.translatable
  - *@versionDate*
- att.typed
  - *@type*

– *@subtype*

*Member of* model.emphLike model.identEquiv

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc category change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage taxonomy

*iso-fs:* fDescr fsDescr

*linking:* ab joinGrp seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* attDef classSpec constraintDecl constraintSpec dataSpec eg elementSpec macroSpec model modelGrp modelSequence moduleSpec paramSpec schemaSpec valDesc valItem

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The *target* and *cRef* attributes are mutually exclusive.

*Example*

```
We may define <term  xml:id="tdpv"  rend="sc">discoursal point of view</term>
as
<gloss  target="#tdpv">the relationship, expressed
through discourse structure, between the implied author or some other
addresser, and the
fiction.</gloss>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element gloss
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cReferencing.attributes,
   att.cmc.attributes,
   att.declaring.attributes,
   att.pointing.attributes,
   att.translatable.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

## <glyph>

(character glyph) provides descriptive information about a character glyph. [*5.2. Markup Constructs for Representation of Characters and Glyphs*]

*Module* gaiji — Characters, Glyphs, and Writing Modes

*Attributes*

- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Contained by* gaiji: charDecl
*May contain*
*core:* binaryObject desc graphic media note noteGrp
*figures:* figure formula
*gaiji:* localProp mapping unicodeProp unihanProp
*Example*

```
<glyph xml:id="rstroke">
  <localProp name="Name"
    value="LATIN SMALL LETTER R WITH A FUNNY STROKE"/>
  <localProp name="entity" value="rstroke"/>
  <figure>
    <graphic url="glyph-rstroke.png"/>
  </figure>
</glyph>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <elementRef key="unicodeProp"/>
  <elementRef key="unihanProp"/>
  <elementRef key="localProp"/>
  <elementRef key="mapping"/>
  <elementRef key="figure"/>
  <classRef key="model.graphicLike"/>
  <classRef key="model.noteLike"/>
  <classRef key="model.descLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element glyph
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      unicodeProp    | unihanProp    | localProp    | mapping    | figure    | model.graph
}
```

---

**<gram>** (grammatical information) within an entry in a dictionary or a terminological data file, contains grammatical information relating to a term, word, or form. [*10.3.2. Grammatical Information*]

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.lexicographic.normalized
    * *@norm*
    * *@orig*

- att.typed
  - ~~type~~
  - @subtype

`@type` classifies grammatical information provided by <gram> according to some convenient and shared typology, ideally one defined in an external reference taxonomy, such as the CLARIN Concept Registry.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **pos** (part of speech) any of the word classes to which a word may be assigned in a given language, based on form, meaning, or a combination of features, e.g. noun, verb, adjective, etc.
**gen** (gender) formal classification by which nouns and pronouns, and often

accompanying modifiers, are grouped and inflected, or changed in form, so as to control certain syntactic relationships

**num** (number) grammatical number, e.g. singular, plural, dual, …

**animate** animate or inanimate

**proper** proper noun or common noun

*Note* The list of values for the *type* attribute should best come from a standardized set of container data categories.

*Member of* model.morphLike

*Contained by*

*dictionaries:* dictScrap entryFree form gramGrp

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<entry>
  <form>
    <orth>pamplemousse</orth>
  </form>
  <gramGrp>
    <gram type="pos">noun</gram>
    <gram type="gen">masculine</gram>
```

```
      </gramGrp>
    </entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element gram
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.paraContent}
```

---

## <gramGrp>

<gramGrp> (grammatical information group) groups morpho-syntactic information about a lexical item, e.g. <pos>, <gen>, <number>, <case>, or <iType> (inflectional class). [*10.3.2. Grammatical Information*]

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.lexicographic.normalized
    * *@norm*
    * *@orig*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.entryPart model.entryPart.top model.lexicalRefinement

*Contained by*

*core:* cit

*dictionaries:* dictScrap entry entryFree etym form gramGrp hom re sense

*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index label lb list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* case colloc gen gram gramGrp iType lang lbl mood number oRef pRef per pos subc tns usg

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<entry>
  <form>
    <orth>luire</orth>
  </form>
  <gramGrp>
    <pos>verb</pos>
    <subc>intransitive</subc>
  </gramGrp>
</entry>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.gramPart"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element gramGrp
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.typed.attributes,
   (
      text
    | model.gLike    | model.phrase    | model.inter    | model.gramPart    | model.global
}
```

---

**<graph>** (graph) encodes a graph, which is a collection of nodes, and arcs which connect the nodes. [*20.1. Graphs and Digraphs*]

*Module* nets — Graphs, Networks, and Trees

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

  * *@cert*
  * *@resp*
- att.global.source
  * *@source*
- att.typed
- ~~type~~
- @subtype

`@type` describes the type of graph.
*Derived from* att.typed
*Status* Recommended
*Datatype* teidata.enumerated
*Suggested values include:* **undirected** undirected graph
**directed** directed graph
**transitionNetwork** a directed graph with distinguished initial and final nodes
**transducer** a transition network with up to two labels on each arc
*Note* If *type* is specified as undirected, then the distinction between the *to* and *from* attributes of the <arc> tag is neutralized. Also, the *adj* attribute, rather than the *adjFrom* and *adjTo* attributes, should be used to encode pointers to the ends of the arcs. If *type* is specified as directed (or any other value which implies directionality), then the *adjFrom* and *adjTo* attributes should be used, instead of the *adj* attribute.

`@order` states the order of the graph, i.e., the number of its nodes.
*Status* Optional
*Datatype* teidata.count

`@size` states the size of the graph, i.e., the number of its arcs.
*Status* Optional
*Datatype* teidata.count

*Member of* model.divPart model.standOffPart
*Contained by*
*core:* item note q quote said stage textLang
*drama:* castList epilogue performance prologue set view
*figures:* cell figure
*header:* change handNote licence scriptNote
*linking:* standOff
*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote
*namesdates:* occupation
*tagdocs:* specGrp
*textcrit:* lem rdg
*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript
*transcr:* metamark
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* cb ellipsis gap gb index label lb milestone note noteGrp pb
*figures:* figure notatedMusic
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*nets:* arc node

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* One or more nodes and zero or more arcs in any order.

*Example*

```
<graph xml:id="cug1" type="undirected"
 order="5" size="4"
 rend="LABEL-PLACE bottom center NODE-FRAME none ARC solid line">
 <label>Airline Connections in Southwestern USA</label>
 <node xml:id="lax" degree="2">
  <label>LAX</label>
 </node>
 <node xml:id="lvg" degree="2">
  <label>LVG</label>
 </node>
 <node xml:id="phx" degree="3">
  <label>PHX</label>
 </node>
 <node xml:id="tus" degree="1">
  <label>TUS</label>
 </node>
 <node xml:id="cib" degree="0">
  <label>CIB</label>
 </node>
 <arc from="#lax" to="#lvg"/>
 <arc from="#lax" to="#phx"/>
 <arc from="#lvg" to="#phx"/>
 <arc from="#phx" to="#tus"/>
</graph>
```

*Content model*

```
<content>
 <sequence>
  <sequence minOccurs="0">
   <elementRef key="label"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <alternate>
   <sequence>
    <sequence minOccurs="1"
     maxOccurs="unbounded">
     <elementRef key="node"/>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
    <sequence minOccurs="0"
     maxOccurs="unbounded">
     <elementRef key="arc"/>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
   </sequence>
   <sequence>
    <sequence minOccurs="1"
```

```
     maxOccurs="unbounded">
     <elementRef key="arc"/>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
    <sequence minOccurs="1"
     maxOccurs="unbounded">
     <elementRef key="node"/>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
   </sequence>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element graph
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type
   {
      "undirected" | "directed" | "transitionNetwork" | "transducer"
   }?,
   attribute order { text }?,
   attribute size { text }?,
   (
      ( ( label, model.global* )? ),
      (
         ( ( ( node, model.global* )+ ), ( ( arc, model.global* )* ) )
       | ( ( ( arc, model.global* )+ ), ( ( node, model.global* )+ ) )
      )
   )
}
```

---

**<graphic>** (graphic) indicates the location of a graphic or illustration, either forming part of a text, or providing an image of it. [*3.10. Graphics and Other Non-textual Components 12.1. Digital Facsimiles*]

*Module* core — Elements Available in All TEI Documents

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.declaring
  - *@decls*
- att.media
  - *@width*
  - *@height*
  - *@scale*

- att.internetMedia
  - *@mimeType*
- att.resourced
  - *@url*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.graphicLike model.titlepagePart

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figure formula notatedMusic table

*gaiji:* char glyph

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material msItem musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePage titlePart trailer

*transcr:* damage facsimile fw metamark mod restore retrace secl sourceDoc supplied surface surplus zone

*verse:* rhyme

*May contain*

*core:* desc

*Note* The *mimeType* attribute should be used to supply the MIME media type of the image specified by the *url* attribute.

Within the body of a text, a <graphic> element indicates the presence of a graphic component in the source itself. Within the context of a <facsimile> or <sourceDoc> element, however, a <graphic> element provides an additional digital representation of some part of the source being encoded.

*Example*

```
<figure>
  <graphic  url="fig1.png"/>
  <head>Figure One: The View from the Bridge</head>
  <figDesc>A Whistleresque view showing four or five sailing boats in the
foreground, and a
    series of buoys strung out between them.</figDesc>
</figure>
```

*Example*

```
<facsimile>
  <surfaceGrp  n="leaf1">
    <surface>
      <graphic  url="page1.png"/>
    </surface>
```

```
      <surface>
        <graphic url="page2-highRes.png"/>
        <graphic url="page2-lowRes.png"/>
      </surface>
    </surfaceGrp>
  </facsimile>
```

*Example*

```
  <facsimile>
    <surfaceGrp n="leaf1" xml:id="spi001">
      <surface xml:id="spi001r">
        <graphic type="normal"
          subtype="thumbnail" url="spi/thumb/001r.jpg"/>
        <graphic type="normal" subtype="low-res"
          url="spi/normal/lowRes/001r.jpg"/>
        <graphic type="normal"
          subtype="high-res" url="spi/normal/highRes/001r.jpg"/>
        <graphic type="high-contrast"
          subtype="low-res" url="spi/contrast/lowRes/001r.jpg"/>
        <graphic type="high-contrast"
          subtype="high-res" url="spi/contrast/highRes/001r.jpg"/>
      </surface>
      <surface xml:id="spi001v">
        <graphic type="normal"
          subtype="thumbnail" url="spi/thumb/001v.jpg"/>
        <graphic type="normal" subtype="low-res"
          url="spi/normal/lowRes/001v.jpg"/>
        <graphic type="normal"
          subtype="high-res" url="spi/normal/highRes/001v.jpg"/>
        <graphic type="high-contrast"
          subtype="low-res" url="spi/contrast/lowRes/001v.jpg"/>
        <graphic type="high-contrast"
          subtype="high-res" url="spi/contrast/highRes/001v.jpg"/>
        <zone xml:id="spi001v_detail01">
          <graphic type="normal"
            subtype="thumbnail" url="spi/thumb/001v-detail01.jpg"/>
          <graphic type="normal"
            subtype="low-res"
            url="spi/normal/lowRes/001v-detail01.jpg"/>
          <graphic type="normal"
            subtype="high-res"
            url="spi/normal/highRes/001v-detail01.jpg"/>
          <graphic type="high-contrast"
            subtype="low-res"
            url="spi/contrast/lowRes/001v-detail01.jpg"/>
          <graphic type="high-contrast"
            subtype="high-res"
            url="spi/contrast/highRes/001v-detail01.jpg"/>
        </zone>
      </surface>
    </surfaceGrp>
  </facsimile>
```

*Content model*

```
<content>
 <classRef key="model.descLike"
```

```
  minOccurs="0" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element graphic
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declaring.attributes,
   att.media.attributes,
   att.internetMedia.attributes,
   att.resourced.attributes,
   att.typed.attributes,
   model.descLike*
}
```

---

**<group>** (group) contains the body of a composite text, grouping together a sequence of distinct texts (or groups of such texts) which are regarded as a unit for some purpose, for example the collected works of an author, a sequence of prose essays, etc. [*4. Default Text Structure 4.3.1. Grouped Texts 16.1. Varieties of Composite Text*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

    * *@prev*
    * *@exclude*
    * *@select*
  – att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  – att.global.responsibility
    * *@cert*
    * *@resp*
  – att.global.source
    * *@source*
- att.declaring
  – *@decls*
- att.typed
  – *@type*
  – *@subtype*

*Contained by* *textstructure:* floatingText group text
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* cb ellipsis gap gb head index lb meeting milestone note noteGrp pb
*figures:* figure notatedMusic
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*spoken:* incident kinesic pause shift vocal writing
*textcrit:* app witDetail
*textstructure:* argument byline closer dateline docAuthor docDate epigraph group opener postscript salute signed text trailer
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
*Example*

```
<text>
<!-- Section on Alexander Pope starts -->
  <front>
<!-- biographical notice by editor -->
  </front>
  <group>
    <text>
<!-- first poem -->
    </text>
    <text>
<!-- second poem -->
    </text>
  </group>
</text>
<!-- end of Pope section-->
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence>
   <alternate>
    <elementRef key="text"/>
    <elementRef key="group"/>
   </alternate>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="text"/>
    <elementRef key="group"/>
    <classRef key="model.global"/>
   </alternate>
  </sequence>
  <classRef key="model.divBottom"
   minOccurs="0" maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element group
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.typed.attributes,
   (
      ( model.divTop | model.global )*,
      ( ( text | group ), ( text | group | model.global )* ),
      model.divBottom*
   )
}
```

---

## <handDesc>

**<handDesc>** (description of hands) contains a description of all the different hands used in a manuscript or other object. [*11.7.2. Writing, Decoration, and Other Notations*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@hands` (hands) specifies the number of distinct hands identified within the manuscript.

*Status* Optional

*Datatype* teidata.count

*Member of* model.physDescPart

*Contained by*

*msdescription:* physDesc

*May contain*

*core:* p

*header:* handNote

*linking:* ab

*msdescription:* summary

*Example*

```
<handDesc>
  <handNote scope="major">Written throughout in <term>angelicana
formata</term>.</handNote>
</handDesc>
```

*Example*

```
<handDesc  hands="2">
  <p>The manuscript is written in two contemporary hands, otherwise
     unknown, but clearly those of practised scribes. Hand I writes
     ff. 1r-22v and hand II ff. 23 and 24. Some scholars, notably
     Verner Dahlerup and Hreinn Benediktsson, have argued for a third hand
     on f. 24, but the evidence for this is insubstantial.</p>
</handDesc>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="summary" minOccurs="0"/>
   <elementRef key="handNote" minOccurs="1"
    maxOccurs="unbounded"/>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element handDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute hands { text }?,
   ( model.pLike+ | ( summary?, handNote+ ) )
}
```

---

## <handNote>

<handNote> (note on hand) describes a particular style or hand distinguished within a manuscript. [*11.7.2. Writing, Decoration, and Other Notations*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*
- att.handFeatures
  - *@scribe*
  - *@scribeRef*
  - *@script*
  - *@scriptRef*
  - *@medium*
- att.scope
  - *@scope*

*Contained by* *msDescription:* handDesc
*transcr:* handNotes
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound spGrp tech view
*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<handNotes>
  <handNote  xml:id="yien.major_hand"
    scope="major"  medium="brown_ink"  scribe="author"
    scribeRef="#KL">Handwriting of the copyist Karol Libelt, brown ink, in
German cursive script.</handNote>
  <handNote  xml:id="yien.pencil"
    scope="minor"  medium="pencil"  scribe="annotator">A note, pencil, unknown
author.</handNote>
  <handNote  xml:id="yien.pencil2"
    scope="minor"  medium="pencil"  scribe="archivist">Writer of the title page
Józef Łepkowski, pencil.</handNote>
  <handNote  xml:id="yien.pencil3"
    scope="minor"  medium="pencil"  scribe="archivist2"
    scribeRef="#BJK">Note on the title page, pencil, by Józef
Korzeniowski.</handNote>
</handNotes>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element handNote
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
```

```
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.handFeatures.attributes,
   att.scope.attributes,
   macro.specialPara}
```

---

**<handNotes>** contains one or more <handNote> elements documenting the different hands identified within the source texts. [*12.3.2.1. Document Hands*]

*Module* transcr — Representation of Primary Sources

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Member of* model.profileDescPart

*Contained by*

*header:* profileDesc
*May contain*
*header:* handNote
*Example*

```
<handNotes>
  <handNote  xml:id="H1"  script="copperplate"
    medium="brown-ink">Carefully written with regular descenders</handNote>
  <handNote  xml:id="H2"  script="print"
    medium="pencil">Unschooled scrawl</handNote>
</handNotes>
```

*Content model*

```
<content>
 <elementRef key="handNote" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element handNotes
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   handNote+
}
```

---

## <handShift>

(handwriting shift) marks the beginning of a sequence of text written in a new hand, or the beginning of a scribal stint. [*12.3.2.1. Document Hands*]

*Module* transcr — Representation of Primary Sources
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  * *@ana*
- att.global.change
  * *@change*
- att.global.facs
  * *@facs*

- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.handFeatures
  - *@scribe*
  - *@scribeRef*
  - *@script*
  - *@scriptRef*
  - *@medium*

- att.scope
  - *@scope*

`@new` indicates a <handNote> element describing the hand concerned.

*Status* Recommended

*Datatype* teidata.pointer

*Note* This attribute serves the same function as the *hand* attribute provided for those elements which are members of the `att.transcriptional` class. It may be renamed at a subsequent major release.

*Member of* model.linePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain* Empty element

*Note* The <handShift> element may be used either to denote a shift in the document hand (as from one scribe to another, on one writing style to another). Or, it may indicate a shift within a document hand, as a change of writing style, character or ink. Like other milestone elements, it should appear at the point of transition from some other state to the state which it describes.

*Example*

```
<l>When wolde the cat dwelle in his ynne</l>
<handShift  medium="greenish-ink"/>
<l>And if the cattes skynne be slyk <handShift  medium="black-ink"/> and
gaye</l>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element handShift
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.handFeatures.attributes,
   att.scope.attributes,
   attribute new { text }?,
   empty
}
```

---

## <head>

(heading) contains any type of heading, for example the title of a section, or the heading of a list, glossary, manuscript description, etc. [*4.2.1. Headings and Trailers*]

*Module* core — Elements Available in All TEI Documents

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.placement
  - *@place*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*

*Member of* model.headLike model.pLike.front

*Contained by*

*cmc:* post

*core:* divGen lg list listBibl

*drama:* castGroup castList epilogue performance prologue set spGrp

*figures:* figure table

*linking:* listAnnotation

*msdescription:* msDesc msFrag msPart

*namesdates:* climate event listEvent listNym listObject listOrg listPerson listPlace listRelation object org place population state terrain trait

*textcrit:* listApp listWit

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 front group postscript

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The <head> element is used for headings at all levels; software which treats (e.g.) chapter headings, section headings, and list titles differently must determine the proper processing of a <head> element based on its structural position. A <head> occurring as the first element of a list is the title of that list; one occurring as the first element of a <div1> is the title of that chapter or section.

*Example* The most common use for the <head> element is to mark the headings of sections. In older writings, the headings or *incipits* may be rather longer than usual in modern works. If a section has an explicit ending as well as a heading, it should be marked as a <trailer>, as in this example:

```
<div1  n="I"  type="book">
  <head>In the name of Christ here begins the first book of the ecclesiastical
history of
    Georgius Florentinus, known as Gregory, Bishop of Tours.</head>
```

```
  <div2  type="section">
    <head>In the name of Christ here begins Book I of the history.</head>
    <p>Proposing as I do ...</p>
    <p>From the Passion of our Lord until the death of Saint Martin four
hundred and twelve
        years passed.</p>
    <trailer>Here ends the first Book, which covers five thousand, five hundred
and ninety-six
        years from the beginning of the world down to the death of Saint
Martin.</trailer>
  </div2>
</div1>
```

*Example* When headings are not inline with the running text (see e.g. the heading ”Secunda conclusio”) they might however be encoded as if. The actual placement in the source document can be captured with the *place* attribute.

```
<div  type="subsection">
  <head  place="margin">Secunda conclusio</head>
  <p>
    <lb  n="1251"/>
    <hi  rend="large">Potencia: habitus: et actus: recipiunt speciem ab
obiectis<supplied>.</supplied>
    </hi>
    <lb  n="1252"/>Probatur sic. Omne importans necessariam habitudinem ad
proprium
     [...]
  </p>
</div>
```

*Example* The <head> element is also used to mark headings of other units, such as lists:

```
With a few exceptions, connectives are equally
useful in all kinds of discourse: description, narration, exposition, argument.
<list  rend="bulleted">
  <head>Connectives</head>
  <item>above</item>
  <item>accordingly</item>
  <item>across from</item>
  <item>adjacent to</item>
  <item>again</item>
  <item>
<!-- ... -->
  </item>
</list>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <elementRef key="lg"/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.lLike"/>
  <classRef key="model.global"/>
```

```
 </alternate>
</content>
```

*Schema Declaration*

```
element head
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.placement.attributes,
   att.typed.attributes,
   att.written.attributes,
   (
      text
    | lg    | model.gLike    | model.phrase    | model.inter    | model.lLike    | model.g
}
```

---

## <headItem>

(heading for list items) contains the heading for the item or gloss column in a glossary list or similar structured list. [*3.8. Lists*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*

* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Contained by* list

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The <headItem> element may appear only if each item in the list is preceded by a <label>.

*Example*

```
The simple, straightforward statement of an
idea is preferable to the use of a worn-out expression. <list type="gloss">
  <headLabel rend="smallcaps">TRITE</headLabel>
  <headItem rend="smallcaps">SIMPLE, STRAIGHTFORWARD</headItem>
  <label>bury the hatchet</label>
  <item>stop fighting, make peace</item>
```

```
      <label>at loose ends</label>
      <item>disorganized</item>
      <label>on speaking terms</label>
      <item>friendly</item>
      <label>fair and square</label>
      <item>completely honest</item>
      <label>at death's door</label>
      <item>near death</item>
    </list>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element headItem
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

---

## <headLabel>

<headLabel> (heading for list labels) contains the heading for the label or term column in a glossary list or similar structured list. [*3.8. Lists*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Contained by* list
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit
*dictionaries:* lang oRef pRef
*figures:* figure formula notatedMusic
*gaiji:* g
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef specDesc specList tag val
*textcrit:* app witDetail
*textstructure:* floatingText
*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo
*verse:* caesura rhyme
character data
*Note* The <headLabel> element may appear only if each item in the list is preceded by a <label>.
*Example*

```
The simple, straightforward statement of an
idea is preferable to the use of a worn-out expression. <list  type="gloss">
  <headLabel  rend="smallcaps">TRITE</headLabel>
  <headItem  rend="smallcaps">SIMPLE, STRAIGHTFORWARD</headItem>
  <label>bury the hatchet</label>
  <item>stop fighting, make peace</item>
  <label>at loose ends</label>
  <item>disorganized</item>
  <label>on speaking terms</label>
  <item>friendly</item>
  <label>fair and square</label>
  <item>completely honest</item>
  <label>at death's door</label>
  <item>near death</item>
</list>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element headLabel
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

---

## <height>

(height) contains a measurement measured along the axis at a right angle to the bottom of the object. [*11.3.4. Dimensions*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

- att.ranging
  - *@atLeast*
  - *@atMost*
  - *@min*
  - *@max*
  - *@confidence*

*Member of* model.dimLike model.measureLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure measureGrp meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote dimensions explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*gaiji:* g

character data

*Example*

```
<height unit="in" quantity="7"/>
```

*Content model* `<content> <macroRef key="macro.xtext"/></content>`

*Schema Declaration*

```
element height
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   macro.xtext}
```

---

## <heraldry>

(heraldry) contains a heraldic formula or phrase, typically found as part of a blazon, coat of arms, etc. [*11.3.8. Heraldry*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<p>Ownership stamp (xvii cent.) on i recto with the arms
<heraldry>A bull passant within a bordure bezanty,
    in chief a crescent for difference</heraldry> [Cole],
crest, and the legend <q>Cole Deum</q>.</p>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element heraldry
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

---

**<hi>** (highlighted) marks a word or phrase as graphically distinct from the surrounding text, for reasons concerning which no claim is made. [*3.3.2.2. Emphatic Words and Phrases 3.3.2. Emphasis, Foreign Words, and Unusual Language*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

    * *@rend*
    * *@style*
    * *@rendition*
    – att.global.responsibility
    * *@cert*
    * *@resp*
    – att.global.source
    * *@source*
  - att.cmc
    – *@generatedBy*
  - att.written
    – *@hand*

*Member of* model.hiLike

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc formula

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout locus material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<hi  rend="gothic">And this Indenture further witnesseth</hi>
that the said <hi  rend="italic">Walter Shandy</hi>, merchant,
in consideration of the said intended marriage ...
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element hi
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
```

```
   att.written.attributes,
   macro.paraContent}
```

---

## <history>

(history) groups elements describing the full history of a manuscript, manuscript part, or other object. [*11.8. History*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by* *msdescription:* msDesc msFrag msPart

*namesdates:* object

*May contain*

*core:* p

*linking:* ab
*msdescription:* acquisition origin provenance summary
*Example*

```
<history>
  <origin>
    <p>Written in Durham during the mid twelfth
        century.</p>
  </origin>
  <provenance>
    <p>Recorded in two medieval
        catalogues of the books belonging to Durham Priory, made in 1391 and
        1405.</p>
  </provenance>
  <provenance>
    <p>Given to W. Olleyf by William Ebchester, Prior (1446-56)
        and later belonged to Henry Dalton, Prior of Holy Island (Lindisfarne)
        according to inscriptions on ff. 4v and 5.</p>
  </provenance>
  <acquisition>
    <p>Presented to Trinity College in 1738 by
        Thomas Gale and his son Roger.</p>
  </acquisition>
</history>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="summary" minOccurs="0"/>
   <elementRef key="origin" minOccurs="0"/>
   <elementRef key="provenance"
    minOccurs="0" maxOccurs="unbounded"/>
   <elementRef key="acquisition"
    minOccurs="0"/>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element history
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.pLike+ | ( summary?, origin?, provenance*, acquisition? ) )
}
```

**<hom>** (homograph) groups information relating to one homograph within an entry. [*10.2. The Structure of Dictionary Entries*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*
- att.lexicographic.normalized
  - *@norm*
  - *@orig*

*Member of* model.entryPart
*Contained by*
*core:* cit
*dictionaries:* dictScrap entry entryFree
*namesdates:* nym
*May contain*
*analysis:* interp interpGrp pc span spanGrp
*certainty:* certainty precision respons
*core:* bibl biblStruct cb cit ellipsis gap gb index lb listBibl milestone note noteGrp pb
*dictionaries:* def dictScrap entry etym form gramGrp lbl re sense usg xr
*figures:* figure notatedMusic
*header:* biblFull
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* msDesc
*spoken:* incident kinesic pause shift vocal writing
*textcrit:* app witDetail
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
*Example*

```
<entry>
  <form>
    <orth>bray</orth>
    <pron>breI</pron>
  </form>
  <hom>
    <gramGrp>
      <pos>n</pos>
    </gramGrp>
    <def>cry of an ass; sound of a trumpet.</def>
  </hom>
  <hom>
    <gramGrp>
      <pos>vt</pos>
      <subc>VP2A</subc>
    </gramGrp>
    <def>make a cry or sound of this kind.</def>
  </hom>
</entry>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
```

```
  <elementRef key="sense"/>
  <elementRef key="pc"/>
  <classRef key="model.entryPart.top"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element hom
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   ( sense | pc | model.entryPart.top | model.global )*
}
```

---

## <hyph>

(hyphenation) contains a hyphenated form of a dictionary headword, or hyphenation information in some other form. [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.lexicographic.normalized
    * *@norm*
    * *@orig*
- att.notated
  - *@notation*

*Member of* model.entryPart model.formPart

*Contained by*

*core:* cit

*dictionaries:* dictScrap entryFree form

*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<entry>
  <form>
    <orth>competitor</orth>
    <hyph>com|peti|tor</hyph>
    <pron>k@m"petit@(r)</pron>
  </form>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element hyph
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.notated.attributes,
   macro.paraContent}
```

**<hyphenation>** (hyphenation) summarizes the way in which hyphenation in a source text has been treated in an encoded version of it. [*2.3.3. The Editorial Practices Declaration 16.3.2. Declarable Elements*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable
  - *@default*

`@eol` (end-of-line) indicates whether or not end-of-line hyphenation has been retained in a text.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **all** all end-of-line hyphenation has been retained, even though the lineation of the original may not have been.

**some** end-of-line hyphenation has been retained in some cases.[Default]

**hard** all soft end-of-line hyphenation has been removed: any remaining end-of-line hyphenation should be retained.

**none** all end-of-line hyphenation has been removed: any remaining hyphenation occurred within the line.

*Member of* model.editorialDeclPart

*Contained by*

*header:* editorialDecl

*May contain*

*core:* p

*linking:* ab

*Example*

```
<hyphenation eol="some">
  <p>End-of-line hyphenation silently removed where appropriate</p>
</hyphenation>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:hyphenation"/> </sch:pattern>

*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element hyphenation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   attribute eol { "all" | "some" | "hard" | "none" }?,
   model.pLike+
}
```

---

## <ident>

(identifier) contains an identifier or name for an object of some kind in a formal language. <ident> is used for tokens such as variable names, class names, type names, function names etc. in formal programming languages. [*23.1.1. Phrase Level Terms*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.typed
  - *@type*
  - *@subtype*

*Member of* model.emphLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain* Character data only

*Note* In running prose, this element may be used for any kind of identifier in any formal language. It should not be used for element and attribute names in XML, for which the special elements <gi> and <att> are provided.

*Example*

```
<ident type="ns">http://www.tei-c.org/ns/Examples</ident>
```

*Content model* `<content> <textNode/></content>`

*Schema Declaration*

```
element ident
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   text
}
```

---

## <idno>

(identifier) supplies any form of identifier used to identify some object, such as a bibliographic item, a person, a title, an organization, etc. in a standardized way. [*14.3.1. Basic Principles 2.2.4. Publication, Distribution, Licensing, etc. 2.2.5. The Series Statement 3.12.2.4. Imprint, Size of a Document, and Reprint Information*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

- att.datable.custom
  - *@when-custom*
  - *@notBefore-custom*
  - *@notAfter-custom*
  - *@from-custom*
  - *@to-custom*
  - *@datingPoint*
  - *@datingMethod*
- att.datable.iso
  - *@when-iso*
  - *@notBefore-iso*

  - * *@notAfter-iso*
  - * *@from-iso*
  - * *@to-iso*
  - – att.datable.w3c
  - * *@when*
  - * *@notBefore*
  - * *@notAfter*
  - * *@from*
  - * *@to*
- att.sortable
  - – *@sortKey*
- att.typed
  - – ~~type~~
  - – @subtype

`@type` categorizes the identifier, for example as an ISBN, Social Security number, etc.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **ISBN** International Standard Book Number: a 13- or (if assigned prior to 2007) 10-digit identifying number assigned by the publishing industry to a published book or similar item, registered with the International ISBN Agency.

**ISSN** International Standard Serial Number: an eight-digit number to uniquely identify a serial publication.

**DOI** Digital Object Identifier: a unique string of letters and numbers assigned to an electronic document.

**URI** Uniform Resource Identifier: a string of characters to uniquely identify a resource, following the syntax of RFC 3986.

**VIAF** A data number in the Virtual Internet Authority File assigned to link different names in catalogs around the world for the same entity.

**ESTC** English Short-Title Catalogue number: an identifying number assigned to a document in English printed in the British Isles or North America before 1801.

**OCLC** OCLC control number (record number) for the union catalog record in WorldCat, a union catalog for member libraries in the Online Computer Library Center global cooperative.

*Member of* model.msItemPart model.nameLike model.personPart model.publicationStmtPart.detail

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address analytic author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned monogr name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote idno language licence principal publicationStmt rendition scriptNote seriesStmt sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions altIdentifier catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msIdentifier msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education event eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality nym objectIdentifier objectName occupation offset org orgName persName persPronouns person personGrp persona place placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg moduleSpec valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*gaiji:* g

*header:* idno

character data

*Note* <idno> should be used for labels which identify an object or concept in a formal cataloguing system such as a database or an RDF store, or in a distributed system such as the World Wide Web. Some suggested values for *type* on <idno> are ISBN, ISSN, DOI, and URI.

*Example*

```
<idno type="ISBN">978-1-906964-22-1</idno>
<idno type="ISSN">0143-3385</idno>
<idno type="DOI">10.1000/123</idno>
<idno type="URI">http://www.worldcat.org/oclc/185922478</idno>
<idno type="URI">http://authority.nzetc.org/463/</idno>
<idno type="LT">Thomason Tract E.537(17)</idno>
<idno type="Wing">C695</idno>
<idno type="oldCat">
  <g ref="#sym"/>345
</idno>
```

In the last case, the identifier includes a non-Unicode character which is defined elsewhere by means of a <glyph> or <char> element referenced here as `#sym`.

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <elementRef key="idno"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element idno
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.sortable.attributes,
   att.typed.attribute.subtype,
   attribute type
   {
      "ISBN" | "ISSN" | "DOI" | "URI" | "VIAF" | "ESTC" | "OCLC"
   }?,
   ( text | model.gLike | idno )*
}
```

---

**<if>** defines a conditional default value for a feature; the condition is specified as a feature structure, and is met if it subsumes the feature structure in the text for which a default value is sought. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Contained by:* *iso-fs:* vDefault
*May contain*
*iso-fs:* binary default f fs numeric string symbol then vAlt vColl vLabel vMerge vNot
*Note* May contain a feature structure, followed by a feature value; the two are separated by a <then> element.
*Example*

```
<vDefault>
  <if>
    <fs>
      <f  name="VFORM">
        <symbol  value="INF"/>
      </f>
      <f  name="SUBJ">
        <binary  value="true"/>
      </f>
    </fs>
    <then/>
    <symbol  value="for"/>
  </if>
</vDefault>
```

*Content model*

```
<content>
 <sequence>
  <alternate>
   <elementRef key="fs"/>
   <elementRef key="f"/>
  </alternate>
  <elementRef key="then"/>
  <classRef key="model.featureVal"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element if
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( ( fs | f ), then, model.featureVal )
}
```

---

**<iff>** (if and only if) separates the condition from the consequence in a <bicond> element. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by:* bicond
*May contain* Empty element
*Note* This element is provided primarily to enhance the human readability of the feature-system declaration.
*Example*

```
<bicond>
  <fs>
    <f  name="FOO">
      <symbol  value="42"/>
    </f>
  </fs>
  <iff/>
  <fs>
    <f  name="BAR">
      <binary  value="true"/>
    </f>
  </fs>
</bicond>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element iff
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   empty
}
```

---

## <imprimatur>

(imprimatur) contains a formal statement authorizing the publication of a work, sometimes required to appear on a title page or its verso. [*4.6. Title Pages*]

*Module* textstructure — Default Text Structure
*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs

    * *@facs*
    – att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
    – att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
    – att.global.responsibility
    * *@cert*
    * *@resp*
    – att.global.source
    * *@source*

*Member of* model.titlepagePart

*Contained by*

*msdescription:* msItem

*textcrit:* lem rdg

*textstructure:* titlePage

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<imprimatur>Licensed and entred acording to Order.</imprimatur>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element imprimatur
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.paraContent}
```

---

**<imprint>** groups information relating to the publication or distribution of a bibliographic item. [*3.12.2.4. Imprint, Size of a Document, and Reprint Information*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*

- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Contained by:* monogr
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* biblScope cb date ellipsis gap gb index lb milestone note noteGrp pb pubPlace publisher respStmt time
*figures:* figure notatedMusic
*header:* catRef classCode distributor
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*spoken:* incident kinesic pause shift vocal writing
*textcrit:* app witDetail
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
*Example*

```
<imprint>
  <pubPlace>Oxford</pubPlace>
  <publisher>Clarendon Press</publisher>
  <date>1987</date>
</imprint>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="classCode"/>
   <elementRef key="catRef"/>
  </alternate>
```

```
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <alternate>
    <classRef key="model.imprintPart"/>
    <classRef key="model.dateLike"/>
   </alternate>
   <elementRef key="respStmt" minOccurs="0"
    maxOccurs="unbounded"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element imprint
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      ( classCode | catRef )*,
      ( ( ( model.imprintPart | model.dateLike ), respStmt*, model.global* )+ )
   )
}
```

---

**<incident>** (incident) marks any phenomenon or occurrence, not necessarily vocalized or communicative, for example incidental noises or other events affecting communication. [*8.3.3. Vocal, Kinesic, Incident*]

*Module* spoken — Transcriptions of Speech

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.ascribed
  - *@who*
- att.cmc
  - *@generatedBy*
- att.timed
  - *@start*
  - *@end*
  - att.duration
    * att.duration.iso
      · *@dur-iso*
    * att.duration.w3c
      · *@dur*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.global.spoken

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* annotationBlock u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*core:* desc

*Example*

```
<incident>
  <desc>ceiling collapses</desc>
</incident>
```

*Content model*

```
<content>
 <classRef key="model.descLike"
  minOccurs="0" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element incident
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.attributes,
   att.cmc.attributes,
```

```
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   att.typed.attributes,
   model.descLike*
}
```

---

**<incipit>** contains the *incipit* of a manuscript or similar object item, that is the opening words of the text proper, exclusive of any *rubric* which might precede it, of sufficient length to identify the work uniquely; such incipits were, in former times, frequently used a means of reference to a work, in place of a title. [*11.6.1. The msItem and msItemStruct Elements*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.msExcerpt

- *@defective*

- att.typed
  - *@type*
  - *@subtype*

*Member of* model.biblPart model.msQuoteLike

*Contained by*

*core:* bibl

*msdescription:* msItem msItemStruct

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<incipit>Pater noster qui es in celis</incipit>
<incipit  defective="true">tatem dedit hominibus alleluia.</incipit>
<incipit  type="biblical">Ghif ons huden onse dagelix broet</incipit>
<incipit>O ongehoerde gewerdighe christi</incipit>
<incipit  type="lemma">Firmiter</incipit>
<incipit>Ideo dicit firmiter quia ordo fidei nostre probari non
potest</incipit>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element incipit
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.msExcerpt.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

## <index> (index entry) marks a location to be indexed for whatever purpose. [*3.9.2. Index Entries*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*

    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.spanning
  - *@spanTo*

`@indexName` a single word which follows the rules defining a legal XML name (see `https://www.w3.org/TR/REC-xml/#dt-name`), supplying a name to specify which index (of several) the index entry belongs to.
*Status* Optional
*Datatype* teidata.name
*Note* This attribute makes it possible to create multiple indexes for a text.

*Member of* model.global.meta
*Contained by*
*analysis:* cl m phr s span w
*cmc:* post
*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint index item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit
*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose
*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view
*figures:* cell figure table
*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor
*linking:* ab seg standOff
*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark
*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname
*nets:* graph
*spoken:* u writing
*tagdocs:* eg valDesc
*textcrit:* lem rdg wit witDetail
*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer
*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone
*verse:* metSym rhyme

*May contain*

*core:* index term

*Example*

```
David's other principal backer, Josiah ha-Kohen
<index  indexName="NAMES">
  <term>Josiah ha-Kohen b. Azarya</term>
</index> b. Azarya, son of one of the last gaons of Sura
<index  indexName="PLACES">
  <term>Sura</term>
</index> was David's own first cousin.
```

*Content model*

```
<content>
 <sequence minOccurs="0"
  maxOccurs="unbounded">
  <elementRef key="term"/>
  <elementRef key="index" minOccurs="0"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element index
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.spanning.attributes,
   attribute indexName { text }?,
   ( ( term, index? )* )
}
```

---

## <iNode> (intermediate (or internal) node) represents an intermediate (or internal) node of a tree. [*20.2. Trees*]

*Module* nets — Graphs, Networks, and Trees

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  * *@ana*

- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@value` indicates an intermediate node, which is a feature structure or other analytic element.
*Status* Optional
*Datatype* teidata.pointer

`@children` provides a list of identifiers of the elements which are the children of the intermediate node.
*Status* Required
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

`@parent` provides the identifier of the element which is the parent of this node.
*Status* Optional
*Datatype* teidata.pointer

`@ord` (ordered) indicates whether or not the internal node is ordered.
*Status* Optional
*Datatype* teidata.xTruthValue
*Note* The value true indicates that the children of the intermediate node are ordered, whereas false indicates the are unordered.
Use if and only if *ord* is specified as partial on the <tree> element and the intermediate node has more than one child.

`@follow` provides the identifier of an element which this node follows.
*Status* Optional
*Datatype* teidata.pointer
*Note* If the tree is unordered or partially ordered, this attribute has the property of fixing the relative order of the intermediate node and the element which is the value of the attribute.

`@outDegree` gives the out degree of an intermediate node, the number of its children.

*Status* Optional
*Datatype* teidata.count
*Note* The in degree of an intermediate node is always 1.

*Contained by* tree
*May contain*
*core:* label
*Example*

```
<iNode xml:id="pt1" children="#GD-UP1"
 parent="#GD-VB1" follow="#GD-PN1" outDegree="1">
 <label>PT</label>
</iNode>
```

*Content model*

```
<content>
 <elementRef key="label" minOccurs="0"/>
</content>
```

*Schema Declaration*

```
element iNode
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute value { text }?,
   attribute children { list { + } },
   attribute parent { text }?,
   attribute ord { text }?,
   attribute follow { text }?,
   attribute outDegree { text }?,
   label?
}
```

---

## <institution>

(institution) contains the name of an organization such as a university or library, with which a manuscript or other object is identified, generally its holding institution. [*11.4. The Manuscript Identifier*]

*Module* msdescription — Manuscript Description
*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.naming
  - *@role*
  - *@nymRef*

  - att.canonical
    * *@key*
    * *@ref*

*Contained by:* *msdescription:* altIdentifier msIdentifier
*namesdates:* objectIdentifier
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit
*dictionaries:* lang
*figures:* figure notatedMusic
*header:* idno
*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data

*Example*

```
<msIdentifier>
  <settlement>Oxford</settlement>
  <institution>University of Oxford</institution>
  <repository>Bodleian Library</repository>
  <idno>MS. Bodley 406</idno>
</msIdentifier>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element institution
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   macro.phraseSeq.limited}
```

---

## <interaction>

(interaction) describes the extent, cardinality and nature of any interaction among those producing and experiencing the text, for example in the form of response or interjection, commentary, etc. [*16.2.1. The Text Description*]

*Module* corpus — Language Corpora

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.typed
  - ~~type~~
  - @subtype

`@type` specifies the degree of interaction between active and passive participants in the text.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **none** no interaction of any kind, e.g. a monologue
**partial** some degree of interaction, e.g. a monologue with set responses
**complete** complete interaction, e.g. a face to face conversation
**inapplicable** this parameter is inappropriate or inapplicable in this case

`@active` specifies the number of active participants (or *addressors*) producing parts of the text.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **singular** a single addressor
**plural** many addressors
**corporate** a corporate addressor
**unknown** number of addressors unknown or unspecifiable

`@passive` specifies the number of passive participants (or *addressees*) to whom a text is directed or in whose presence it is created or performed.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **self** text is addressed to the originator e.g. a diary
**single** text is addressed to one other person e.g. a personal letter
**many** text is addressed to a countable number of others e.g. a conversation in which all participants are identified
**group** text is addressed to an undefined but fixed number of participants e.g. a lecture
**world** text is addressed to an undefined and indeterminately large number e.g. a published book

*Member of* model.textDescPart
*Contained by*
*corpus:* textDesc
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit
*dictionaries:* lang
*figures:* figure notatedMusic
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef tag val
*textcrit:* app witDetail
*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin
character data
*Example*

```
<interaction type="complete"
 active="plural" passive="many"/>
```

*Example*

```
<interaction type="none" active="singular"
 passive="group"/>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
```

```
</content>
```

*Schema Declaration*

```
element interaction
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { "none" | "partial" | "complete" | "inapplicable" }?,
   attribute active { "singular" | "plural" | "corporate" | "unknown" }?,
   attribute passive { "self" | "single" | "many" | "group" | "world" }?,
   macro.phraseSeq.limited}
```

---

**<interp>** (interpretation) summarizes a specific interpretative annotation which can be linked to a span of text. [*18.3. Spans and Interpretations*]

*Module* analysis — Simple Analytic Mechanisms

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.interpLike
  - *@type*
  - *@subtype*
  - *@inst*

*Member of* model.global.meta

*Contained by*

*analysis:* cl interpGrp m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*certainty:* certainty precision respons

*core:* desc

*gaiji:* g

character data

*Note* Generally, each <interp> element carries an *xml:id* attribute. This permits the encoder to explicitly associate the interpretation represented by the content of an <interp> with any textual element through its *ana* attribute.

Alternatively (or, in addition) an <interp> may carry an *inst* attribute that points to one or more textual elements to which the analysis represented by the content of the <interp> applies.

*Example*

```
<interp type="structuralunit"
  xml:id="ana_am">aftermath</interp>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.descLike"/>
  <classRef key="model.certLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element interp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.interpLike.attributes,
   ( text | model.gLike | model.descLike | model.certLike )*
}
```

---

**<interpGrp>** (interpretation group) collects together a set of related interpretations which share responsibility or type. [*18.3. Spans and Interpretations*]

*Module* analysis — Simple Analytic Mechanisms

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.interpLike
  - *@type*
  - *@subtype*
  - *@inst*

*Member of* model.global.meta

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*analysis:* interp

*core:* desc

*Note* Any number of <interp> elements.

*Example*

```
<interpGrp  resp="#TMA"
  type="structuralunit">
  <desc>basic structural organization</desc>
  <interp  xml:id="I1">introduction</interp>
  <interp  xml:id="I2">conflict</interp>
  <interp  xml:id="I3">climax</interp>
  <interp  xml:id="I4">revenge</interp>
  <interp  xml:id="I5">reconciliation</interp>
  <interp  xml:id="I6">aftermath</interp>
</interpGrp>
<bibl  xml:id="TMA">
<!-- bibliographic citation for source of this
interpretive framework -->
</bibl>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.descLike"
```

```
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="interp" minOccurs="1"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element interpGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.interpLike.attributes,
   ( model.descLike*, interp+ )
}
```

---

**<interpretation>** (interpretation) describes the scope of any analytic or interpretive information added to the text in addition to the transcription. [*2.3.3. The Editorial Practices Declaration*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.declarable
  - *@default*

*Member of* model.editorialDeclPart

*Contained by*

*header:* editorialDecl

*May contain*

*core:* p

*linking:* ab

*Example*

```
<interpretation>
  <p>The part of speech analysis applied throughout section 4 was added by hand
and has not
    been checked</p>
</interpretation>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:interpretation"/> </sch:pattern>

*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element interpretation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   model.pLike+
}
```

**<item>** (item) contains one component of a list. [*3.8. Lists 2.6. The Revision Description*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.sortable
  - *@sortKey*

*Contained by* *core:* list

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media

mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May contain simple prose or a sequence of chunks.

Whatever string of characters is used to label a list item in the copy text may be used as the value of the global *n* attribute, but it is not required that numbering be recorded explicitly. In ordered lists, the *n* attribute on the <item> element is by definition synonymous with the use of the <label> element to record the enumerator of the list item. In glossary lists, however, the term being defined should be given with the <label> element, not *n*.

*Example*

```
<list  rend="numbered">
  <head>Here begin the chapter headings of Book IV</head>
  <item  n="4.1">The death of Queen Clotild.</item>
  <item  n="4.2">How King Lothar wanted to appropriate one third of the Church
revenues.</item>
  <item  n="4.3">The wives and children of Lothar.</item>
  <item  n="4.4">The Counts of the Bretons.</item>
  <item  n="4.5">Saint Gall the Bishop.</item>
  <item  n="4.6">The priest Cato.</item>
  <item> ...</item>
</list>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element item
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.sortable.attributes,
   macro.specialPara}
```

---

**<iType>** (inflectional class) indicates the inflectional class associated with a lexical item. [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*

  * *@resp*
  - att.global.source
  * *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*
  - att.datcat
  * *@datcat*
  * *@valueDatcat*
  * *@targetDatcat*
  - att.lexicographic.normalized
  * *@norm*
  * *@orig*
- att.typed
  - ~~type~~
  - @subtype

`@type` indicates the type of indicator used to specify the inflection class, when it is necessary to distinguish between the usual abbreviated indications (e.g. *inv*) and other kinds of indicators, such as special codes referring to conjugation patterns, etc.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **abbrev** abbreviated indicator
**verbTable** coded reference to a table of verbs
*Note* This element is synonymous with <gram type='inflectional type'>.

*Member of* model.morphLike
*Contained by*
*dictionaries:* dictScrap entryFree form gramGrp
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound tech view
*figures:* figure formula notatedMusic table
*gaiji:* g
*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May contain character data and phrase-level elements. Typical content will be *invariant*, *n 3* etc.

*Example*

```
<form>
  <orth>horrifier</orth>
  <pron>ORifje</pron>
  <iType  type="vbtable">7</iType>
</form>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element iType
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.paraContent}
```

**<join>** (join) identifies a possibly fragmented segment of text, by pointing at the possibly discontiguous elements which compose it. [*17.7. Aggregation*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*
- att.typed
  - *@type*
  - *@subtype*

`@result` specifies the name of an element which this aggregation may be understood to represent.
*Status* Optional
*Datatype* teidata.name

`@scope` indicates whether the targets to be joined include the entire element indicated (the entire subtree including its root), or just the children of the target (the branches of the subtree).
*Status* Recommended
*Datatype* teidata.enumerated
*Legal values are:* **root** the rooted subtrees indicated by the *target* attribute are joined, each subtree become a child of the virtual element created by the join[Default]
**branches** the children of the subtrees indicated by the *target* attribute become the children of the virtual element (i.e. the roots of the subtrees are discarded)

*Member of* model.global.meta
*Contained by*
*analysis:* cl m phr s span w
*cmc:* post
*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit
*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose
*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view
*figures:* cell figure table
*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor
*linking:* ab joinGrp seg standOff
*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark
*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname
*nets:* graph
*spoken:* u writing
*tagdocs:* eg valDesc
*textcrit:* lem rdg wit witDetail
*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer
*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone
*verse:* metSym rhyme
*May contain*
*certainty:* certainty precision respons
*core:* desc

*Example* The following example is discussed in section *17.7. Aggregation*:

```
<sp>
  <speaker>Hughie</speaker>
  <p>How does it go? <q>
     <l  xml:id="frog_x1">da-da-da</l>
     <l  xml:id="frog_l2">gets a new frog</l>
     <l>...</l>
    </q>
  </p>
</sp>
<sp>
  <speaker>Louie</speaker>
  <p>
    <q>
     <l  xml:id="frog_l1">When the old pond</l>
     <l>...</l>
    </q>
  </p>
</sp>
<sp>
  <speaker>Dewey</speaker>
  <p>
    <q>... <l  xml:id="frog_l3">It's a new pond.</l>
    </q>
  </p>
  <join  target="#frog_l1 #frog_l2 #frog_l3"
    result="lg"  scope="root"/>
</sp>
```

The <join> element here identifies a linegroup (<lg>) comprising the three lines indicated by the *target* attribute. The value root for the *scope* attribute indicates that the resulting virtual element contains the three <l> elements linked to at #frog_l1 #frog_l2 #frog_l3, rather than their character data content.

*Example* In this example, the attribute *scope* is specified with the value of branches to indicate that the virtual list being constructed is to be made by taking the lists indicated by the *target* attribute of the <join> element, discarding the <list> tags which enclose them, and combining the items contained within the lists into a single virtual list:

```
<p>Southern dialect (my own variety, at least) has only <list  xml:id="LP1">
    <item>
      <s>I done gone</s>
    </item>
    <item>
      <s>I done went</s>
    </item>
  </list> whereas Negro Non-Standard basilect has both these and
<list  xml:id="LP2">
    <item>
      <s>I done go</s>
    </item>
  </list>.</p>
<p>White Southern dialect also has <list  xml:id="LP3">
    <item>
      <s>I've done gone</s>
    </item>
    <item>
      <s>I've done went</s>
    </item>
```

```
    </list> which, when they occur in Negro dialect, should probably be
  considered as borrowings from other varieties of English.</p>
  <join  result="list"  xml:id="LST1"
    target="#LP1 #LP2 #LP3"  scope="branches">
    <desc>Sample sentences in Southern speech</desc>
  </join>
```

*Schematron* <sch:rule context="tei:join"> <sch:assert test="contains( normalize-space( @target ),' ')"> You must supply at least two values for @target on <sch:name/> </sch:assert> </sch:rule>

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.descLike"/>
  <classRef key="model.certLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element join
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.pointing.attributes,
   att.typed.attributes,
   attribute result { text }?,
   attribute scope { "root" | "branches" }?,
   ( model.descLike | model.certLike )*
}
```

---

## <joinGrp>

(join group) groups a collection of <join> elements and possibly pointers. [*17.7. Aggregation*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.pointing.group
  - *@domains*
  - *@targFunc*

- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*
- att.typed
  - *@type*
  - *@subtype*

`@result` supplies the default value for the *result* on each <join> included within the group.

*Status* Optional

*Datatype* teidata.name

*Member of* model.global.meta

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note

num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*core:* desc gloss ptr

*linking:* join

*tagdocs:* equiv

*Note* Any number of <join> or <ptr> elements.

*Example*

```
<joinGrp domains="#zuitxt1 #zuitxt2 #zuitxt3"
  result="q">
  <join target="#zuiq1 #zuiq2 #zuiq6"/>
  <join target="#zuiq3 #zuiq4 #zuiq5"/>
</joinGrp>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="equiv"/>
   <elementRef key="gloss"/>
   <classRef key="model.descLike"/>
  </alternate>
```

```
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="join"/>
   <elementRef key="ptr"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element joinGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.pointing.group.attributes,
   att.pointing.attributes,
   att.typed.attributes,
   attribute result { text }?,
   ( ( equiv | gloss | model.descLike )*, ( join | ptr )+ )
}
```

---

**<keywords>** (keywords) contains a list of keywords or phrases identifying the topic or nature of a text. [*2.4.3. The Text Classification*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

`@scheme` identifies the controlled vocabulary within which the set of keywords concerned is defined, for example by a <taxonomy> element, or by some other resource.
*Status* Optional
*Datatype* teidata.pointer

*Contained by:* textClass
*May contain*
*core:* list term
*Note* Each individual keyword (including compound subject headings) should be supplied as a <term> element directly within the <keywords> element. An alternative usage, in which each <term> appears within an <item> inside a <list> is permitted for backwards compatibility, but is deprecated.

If no control list exists for the keywords used, then no value should be supplied for the *scheme* attribute.

*Example*

```
<keywords  scheme="http://classificationweb.net">
  <term>Babbage, Charles</term>
  <term>Mathematicians - Great Britain - Biography</term>
</keywords>
```

*Example*

```
<keywords>
  <term>Fermented beverages</term>
  <term>Central Andes</term>
  <term>Schinus molle</term>
  <term>Molle beer</term>
  <term>Indigenous peoples</term>
  <term>Ethnography</term>
  <term>Archaeology</term>
</keywords>
```

*Content model*

```
<content>
 <alternate>
  <elementRef key="term" minOccurs="1"
```

```
   maxOccurs="unbounded"/>
  <elementRef key="list"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element keywords
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute scheme { text }?,
   ( term+ | list )
}
```

---

**<kinesic>** (kinesic) marks any communicative phenomenon, not necessarily vocalized, for example a gesture, frown, etc. [*8.3.3. Vocal, Kinesic, Incident*]

*Module* spoken — Transcriptions of Speech

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*
- att.ascribed.directed
  - *@toWhom*
- att.ascribed
  - *@who*
- att.cmc
  - *@generatedBy*
- att.timed
  - *@start*
  - *@end*
- att.duration
  - att.duration.iso
    - *@dur-iso*
  - att.duration.w3c
    - *@dur*
- att.typed
  - *@type*
  - *@subtype*

`@iterated` indicates whether or not the phenomenon is repeated.

*Status* Optional

*Datatype* teidata.xTruthValue

*Default* false

*Note* The value true indicates that the kinesic is repeated several times rather than occurring only once.

*Member of* model.global.spoken

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* annotationBlock u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*core:* desc

*Example*

```
<kinesic dur="PT1.5S" iterated="true"
  type="reinforcing">
  <desc>nodding head vigorously</desc>
</kinesic>
```

*Content model*

```
<content>
 <classRef key="model.descLike"
  minOccurs="0" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element kinesic
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.directed.attributes,
```

```
   att.ascribed.attributes,
   att.cmc.attributes,
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   att.typed.attributes,
   attribute iterated { text }?,
   model.descLike*
}
```

---

**<l>** (verse line) contains a single, possibly incomplete, line of verse. [*3.13.1. Core Tags for Verse 3.13. Passages of Verse or Drama 7.2.5. Speech Contents*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

- att.cmc
  - *@generatedBy*
- att.enjamb
  - *@enjamb*
- att.fragmentable
  - *@part*
- att.metrical
  - *@met*
  - *@real*
  - *@rhyme*

*Member of* model.lLike

*Contained by*

*cmc:* post

*core:* add corr del emph head hi item lg note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def gen gram hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figure

*header:* change handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*tagdocs:* specGrp

*textcrit:* lem rdg

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index label lb list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat

geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<l  met="x/x/x/x/x/"  real="/xx/x/x/x/">Shall I compare thee to a summer's
day?</l>
```

*Schematron* <sch:rule context="tei:l"> <sch:report test="ancestor::tei:l[not(.//tei:note//tei:l[. = current()])]">Abstract model violation: Lines may not contain lines or lg elements.</sch:report> </sch:rule>

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element l
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.enjamb.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   ( text | model.gLike | model.phrase | model.inter | model.global )*
}
```

**<label>** (label) contains any label or heading used to identify part of a text, typically but not exclusively in a list or glossary. [*3.8. Lists*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.placement
  - *@place*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*

*Member of* model.labelLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l lg list meeting note orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure notatedMusic

*header:* application change handNote licence rendition scriptNote tagUsage unitDef

*iso-fs:* fDescr fsDescr

*linking:* ab listAnnotation seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* climate event location occupation org place population state terrain trait

*nets:* arc eLeaf eTree graph iNode leaf node root tree triangle

*spoken:* transcriptionDesc u writing

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surface surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example* Labels are commonly used for the headwords in glossary lists; note the use of the global *xml:lang*

attribute to set the default language of the glossary list to Middle English, and identify the glosses and headings as modern English or Latin:

```
<list  type="gloss"  xml:lang="enm">
  <head  xml:lang="en">Vocabulary</head>
  <headLabel  xml:lang="en">Middle English</headLabel>
  <headItem  xml:lang="en">New English</headItem>
  <label>nu</label>
  <item  xml:lang="en">now</item>
  <label>lhude</label>
  <item  xml:lang="en">loudly</item>
  <label>bloweth</label>
  <item  xml:lang="en">blooms</item>
  <label>med</label>
  <item  xml:lang="en">meadow</item>
  <label>wude</label>
  <item  xml:lang="en">wood</item>
  <label>awe</label>
  <item  xml:lang="en">ewe</item>
  <label>lhouth</label>
  <item  xml:lang="en">lows</item>
  <label>sterteth</label>
  <item  xml:lang="en">bounds, frisks (cf. <cit>
     <ref>Chaucer, K.T.644</ref>
     <quote>a courser, <term>sterting</term>as the fyr</quote>
    </cit>
  </item>
  <label>verteth</label>
  <item  xml:lang="la">pedit</item>
  <label>murie</label>
  <item  xml:lang="en">merrily</item>
  <label>swik</label>
  <item  xml:lang="en">cease</item>
  <label>naver</label>
  <item  xml:lang="en">never</item>
</list>
```

*Example* Labels may also be used to record explicitly the numbers or letters which mark list items in ordered lists, as in this extract from Gibbon's *Autobiography*. In this usage the <label> element is synonymous with the *n* attribute on the <item> element:

```
I will add two facts, which have seldom occurred
in the composition of six, or at least of five quartos.
<list  rend="runon"  type="ordered">
  <label>(1)</label>
  <item>My first rough manuscript, without any intermediate copy, has been sent
to the press.</item>
  <label>(2) </label>
  <item>Not a sheet has been seen by any human eyes, excepting those of the
author and the
     printer: the faults and the merits are exclusively my own.</item>
</list>
```

*Example* Labels may also be used for other structured list items, as in this extract from the journal of Edward Gibbon:

```
<list  type="gloss">
  <label>March 1757.</label>
```

```
  <item>I wrote some critical observations upon Plautus.</item>
  <label>March 8th.</label>
  <item>I wrote a long dissertation upon some lines of Virgil.</item>
  <label>June.</label>
  <item>I saw Mademoiselle Curchod — <quote  xml:lang="la">Omnia vincit amor,
et nos cedamus
       amori.</quote>
  </item>
  <label>August.</label>
  <item>I went to Crassy, and staid two days.</item>
</list>
```

Note that the <label> might also appear within the <item> rather than as its sibling. Though syntactically valid, this usage is not recommended TEI practice.

*Example* Labels may also be used to represent a label or heading attached to a paragraph or sequence of paragraphs not treated as a structural division, or to a group of verse lines. Note that, in this case, the <label> element appears *within* the <p> or <lg> element, rather than as a preceding sibling of it.

```
<p>[...]
<lb/>& n’entrer en mauuais & mal-heu-
<lb/>ré meſnage. Or des que le conſente-
<lb/>ment des parties y eſt le mariage eſt
<lb/> arreſté, quoy que de faict il ne ſoit
<label  place="margin">Puiſſance maritale
    entre les Romains.</label>
  <lb/> conſommé. Depuis la conſomma-
<lb/>tion du mariage la femme eſt ſoubs
<lb/> la puiſſance du mary, s’il n’eſt eſcla-
<lb/>ue ou enfant de famille : car en ce
<lb/> cas, la femme, qui a eſpouſé vn en-
<lb/>fant de famille, eſt ſous la puiſſance
[...]</p>
```

In this example the text of the label appears in the right hand margin of the original source, next to the paragraph it describes, but approximately in the middle of it. If so desired the *type* attribute may be used to distinguish different categories of label.

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element label
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.placement.attributes,
```

```
   att.typed.attributes,
   att.written.attributes,
   macro.phraseSeq}
```

---

**<lacunaEnd>** (lacuna end) indicates the end of a lacuna in a mostly complete textual witness. [*13.1.5. Fragmentary Witnesses*]

*Module* textcrit — Critical Apparatus

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.rdgPart
  - *@wit*

*Member of* model.rdgPart

*Contained by*

*textcrit:* lem rdg

*May contain* Empty element

*Example*

```
<rdg  wit="#X">
  <lacunaEnd/>auctorite
</rdg>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element lacunaEnd
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.rdgPart.attributes,
   empty
}
```

---

## <lacunaStart>

<lacunaStart> (lacuna start) indicates the beginning of a lacuna in the text of a mostly complete textual witness. [*13.1.5. Fragmentary Witnesses*]

*Module* textcrit — Critical Apparatus

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*

    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.rdgPart
  - *@wit*

*Member of* model.rdgPart
*Contained by*
*textcrit:* lem rdg
*May contain* Empty element
*Example*

```
<app>
  <lem  wit="#El #Hg">Experience</lem>
  <rdg  wit="#Ha4">Ex<g  ref="#per"/>
    <lacunaStart/>
  </rdg>
</app>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element lacunaStart
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.rdgPart.attributes,
   empty
}
```

---

## <lang>

(language name) contains the name of a language mentioned in etymological or other linguistic discussion. [*10.3.4. Etymological Information*]

*Module* dictionaries — Dictionaries
*Attributes*
- att.global
  - *@xml:id*
  - *@n*

  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.lexicographic.normalized
    * *@norm*
    * *@orig*

*Member of* model.nameLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May contain character data mixed with phrase-level elements.

*Example*

```
<entry>
  <form>
    <orth>publish</orth> ... </form>
  <etym>
    <lang>ME.</lang>
    <mentioned>publisshen</mentioned>,
  <lang>F.</lang>
    <mentioned>publier</mentioned>, <lang>L.</lang>
    <mentioned>publicare,
        publicatum</mentioned>. <xr>See <ref>public</ref>; cf. <ref>2d
-ish</ref>.</xr>
  </etym>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element lang
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   macro.paraContent}
```

**<langKnowledge>** (language knowledge) summarizes the state of a person's linguistic knowledge, either as prose or by a list of <langKnown> elements. [*14.3.2.1. Personal Characteristics*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*
    - *@to-custom*
    - *@datingPoint*

  * *@datingMethod*
  - att.datable.iso
  * *@when-iso*
  * *@notBefore-iso*
  * *@notAfter-iso*
  * *@from-iso*
  * *@to-iso*
  - att.datable.w3c
  * *@when*
  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **listening**
**speaking**
**reading**
**writing**

`@tags` supplies one or more valid language tags for the languages specified.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.language separated by whitespace
*Note* This attribute should be supplied only if the element contains no <langKnown> children. Its values are language 'tags' as defined in RFC 4646 or its successor

*Member of* model.persStateLike
*Contained by*
*namesdates:* person personGrp persona
*May contain*
*certainty:* precision
*core:* p
*linking:* ab
*namesdates:* langKnown
*Example*

```
<langKnowledge tags="en-GB fr">
```

```
    <p>British English and French</p>
  </langKnowledge>
```

*Example*

```
  <langKnowledge>
    <langKnown  tag="en-GB"  level="H">British English</langKnown>
    <langKnown  tag="fr"  level="M">French</langKnown>
  </langKnowledge>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="precision" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate>
   <classRef key="model.pLike"/>
   <elementRef key="langKnown"
    minOccurs="1" maxOccurs="unbounded"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element langKnowledge
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   attribute tags { list { + } }?,
   ( precision*, ( model.pLike | langKnown+ ) )
}
```

---

**<langKnown>** (language known) summarizes the state of a person's linguistic competence, i.e., knowledge of a single language. [*16.2.2. The Participant Description*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*

  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*

    * *@from-iso*
    * *@to-iso*
  – att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  – *@evidence*
  – *@instant*

`@tag` supplies a valid language tag for the language concerned.
*Status* Required
*Datatype* teidata.language
*Note* The value for this attribute should be a language 'tag' as defined in BCP 47.

`@level` a code indicating the person's level of knowledge for this language.
*Status* Optional
*Datatype* teidata.word

*Contained by:* namesdates: langKnowledge
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit
*dictionaries:* lang
*figures:* figure notatedMusic
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef tag val
*textcrit:* app witDetail
*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin
character data

*Example*

```
<langKnown tag="en-GB" level="H">British English</langKnown>
<langKnown tag="fr" level="M">French</langKnown>
```

*Example*

```
<person  sex="m"  role="speaker collaborator">
<!-- other details omitted -->
  <langKnowledge>
    <langKnown  tag="mix">Mixtepec-Mixtec</langKnown>
    <langKnown  tag="en">English</langKnown>
    <langKnown  tag="es">Spanish</langKnown>
  </langKnowledge>
</person>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element langKnown
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   attribute tag { text },
   attribute level { text }?,
   macro.phraseSeq.limited}
```

---

## <language>

(language) characterizes a single language or sublanguage used within a text. [*2.4.2. Language Usage*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.scope
  - *@scope*

`@ident` (identifier) Supplies a language code constructed as defined in BCP 47 which is used to identify the language documented by this element, and which may be referenced by the global *xml:lang* attribute.
*Status* Required
*Datatype* teidata.language

`@usage` specifies the approximate percentage of the text which uses this language.
*Status* Optional
*Datatype* nonNegativeInteger

*Contained by:* header: langUsage
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit
*dictionaries:* lang
*figures:* figure notatedMusic
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat

geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data

*Note* Particularly for sublanguages, an informal prose characterization should be supplied as content for the element.

*Example*

```
<langUsage>
  <language ident="en-US" usage="75">modern American English</language>
  <language ident="az-Arab" usage="20">Azerbaijani in Arabic script</language>
  <language ident="x-lap" usage="05">Pig Latin</language>
</langUsage>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element language
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.scope.attributes,
   attribute ident { text },
   attribute usage { text }?,
   macro.phraseSeq.limited}
```

---

## <langUsage>

(language usage) describes the languages, sublanguages, registers, dialects, etc. represented within a text. [*2.4.2. Language Usage 2.4. The Profile Description 16.3.2. Declarable Elements*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.declarable
  - *@default*

*Member of* model.profileDescPart
*Contained by*
*header:* profileDesc
*May contain*
*core:* p
*header:* language
*linking:* ab
*Example*

```
<langUsage>
  <language ident="fr-CA" usage="60">Québecois</language>
  <language ident="en-CA" usage="20">Canadian business English</language>
  <language ident="en-GB" usage="20">British English</language>
</langUsage>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:langUsage"/> </sch:pattern>
*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
```

```
   maxOccurs="unbounded"/>
  <elementRef key="language" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element langUsage
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   ( model.pLike+ | language+ )
}
```

---

**<layout>** (layout) describes how text is laid out on the page or surface of the object, including information about any ruling, pricking, or other evidence of page-preparation techniques. [*11.7.2. Writing, Decoration, and Other Notations*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@columns` (columns) specifies the number of columns per page.

*Status* Optional

*Datatype* 1–2 occurrences of teidata.count separated by whitespace

*Note* If a single number is given, all pages referenced have this number of columns. If two numbers are given, the number of columns per page varies between the values supplied. Where *columns* is omitted the number is assumed to be 1.

Columns may be independent of page orientation or reading direction, and a single textual *stream* may have one or more columns.

`@streams` (textual streams) indicates the number of streams per page, each of which contains an independent textual stream.

*Status* Optional

*Datatype* 1–2 occurrences of teidata.count separated by whitespace

*Note* If a single number is given, all pages referenced have this number of textual streams. If two numbers are given, the number of textual streams per page varies between the values supplied. Where *streams* is omitted the number is assumed to be 1 and unless specified elsewhere the script orientation of the source is identical to that used in the TEI document.

`@ruledLines` (ruled lines) specifies the number of ruled lines per column.

*Status* Optional

*Datatype* 1–2 occurrences of teidata.count separated by whitespace

*Note* If a single number is given, all columns have this number of ruled lines. If two numbers are given, the number of ruled lines per column varies between the values supplied.

`@writtenLines` (written lines) specifies the number of written lines per column.

*Status* Optional

*Datatype* 1–2 occurrences of teidata.count separated by whitespace

*Note* If a single number is given, all columns have this number of written lines. If two numbers are given, the number of written lines per column varies between the values supplied.

*Contained by:* *desc* layoutDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<layout columns="1" ruledLines="25 32">Most pages have between 25 and 32 long
lines ruled in lead.</layout>
```

*Example*

```
<layout columns="2" ruledLines="42">
  <p>2 columns of 42 lines ruled in ink, with central rule
     between the columns.</p>
</layout>
```

*Example*

```
<layout columns="1 2" writtenLines="40 50">
  <p>Some pages have 2 columns, with central rule
     between the columns; each column with between 40 and 50 lines of
writing.</p>
</layout>
```

*Example*

```
<layout streams="3" columns="3"/>
<!-- Further down in document body -->
<div type="page">
  <ab>
    <pb/>
     □□□<cb type="top-stream"/>
     □□□<cb type="mid-stream"/>
     □□□<cb type="bottom-stream"/>
```

```
    <!-- cb here for demo purposes -->
      </ab>
    </div>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element layout
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute columns { list { ? } }?,
   attribute streams { list { ? } }?,
   attribute ruledLines { list { ? } }?,
   attribute writtenLines { list { ? } }?,
   macro.specialPara}
```

---

**<layoutDesc>** (layout description) collects the set of layout descriptions applicable to a manuscript or other object. [*11.7.2. Writing, Decoration, and Other Notations*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

*Contained by:* msdescription: objectDesc
*May contain*
*core:* p
*linking:* ab
*msdescription:* layout summary
*Example*

```
<layoutDesc>
  <p>Most pages have between 25 and 32 long lines ruled in lead.</p>
</layoutDesc>
```

*Example*

```
<layoutDesc>
  <layout  columns="2"  ruledLines="42">
    <p>
      <locus  from="f12r"  to="f15v"/>
         2 columns of 42 lines pricked and ruled in ink, with
         central rule between the columns.</p>
  </layout>
  <layout  columns="3">
    <p>
      <locus  from="f16"/>Prickings for three columns are visible.</p>
  </layout>
</layoutDesc>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="summary" minOccurs="0"/>
   <elementRef key="layout" minOccurs="1"
    maxOccurs="unbounded"/>
  </sequence>
```

```
 </alternate>
</content>
```

*Schema Declaration*

```
element layoutDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.pLike+ | ( summary?, layout+ ) )
}
```

---

**<lb>** (line beginning) marks the beginning of a topographic line in some edition or version of a text. [*3.11.3. Milestone Elements 7.2.5. Speech Contents*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.breaking
  - *@break*
- att.cmc
  - *@generatedBy*
- att.edition
  - *@ed*
  - *@edRef*
- att.spanning
  - *@spanTo*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.milestoneLike

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list listBibl measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc subst supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Note* By convention, <lb> elements should appear at the point in the text where a new line starts. The *n* attribute, if used, indicates the number or other value associated with the text between this point and the next <lb> element, typically the sequence number of the line within the page, or other appropriate unit. This element is intended to be used for marking the beginning of each new topographic line on a manuscript or printed page, at the point where it occurs; it should not be used to tag structural units such as lines of verse (for which the <l> element is available) except in circumstances where structural units cannot otherwise be marked.

The *type* attribute may be used to characterize the line beginning in any respect. The more specialized attributes *break*, *ed*, or *edRef* should be preferred when the intent is to indicate whether or not the beginning of the new topographic line is word-breaking, or to note the source from which it derives.

*Example* This example shows the encoding of the beginning of each new topographic line within a metrical line, indicating where it occurs in both the 1667 and 1674 editions:

```
<l>Of Mans First Disobedience,<lb ed="1674"/> and<lb ed="1667"/> the
Fruit</l>
<l>Of that Forbidden Tree, whose<lb ed="1667 1674"/> mortal tast</l>
<l>Brought Death into the World,<lb ed="1667"/> and all<lb ed="1674"/> our
woe,</l>
```

*Example* This example shows the encoding of the beginning of a new topographical line as a means of preserving the visual appearance of a title page. The *break* attribute is used to show that the beginning of the new line does not (as elsewhere) mark the start of a new word.

```
<titlePart>
  <lb/>With Additions, ne-<lb break="no"/>ver before Printed.
</titlePart>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element lb
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.breaking.attributes,
   att.cmc.attributes,
   att.edition.attributes,
   att.spanning.attributes,
   att.typed.attributes,
   empty
}
```

**<lbl>** (label) contains a label for a form, example, translation, or other piece of information, e.g. abbreviation for, contraction of, literally, approximately, synonyms:, etc. [*10.3.1. Information on Written and Spoken Forms 10.3.3.2. Translation Equivalents 10.3.5.3. Cross-References to Other Entries*]

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*
- att.lexicographic.normalized
  - *@norm*
  - *@orig*
- att.typed
  - ~~type~~
  - @subtype

`@type` classifies the label using any convenient typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated

*Member of* model.entryPart model.entryPart.top model.lexicalRefinement
*Contained by*
*core:* cit
*dictionaries:* dictScrap entry entryFree etym form gramGrp hom re sense xr
*namesdates:* nym
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound tech view
*figures:* figure formula notatedMusic table
*gaiji:* g
*header:* biblFull idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val
*textcrit:* app listApp listWit witDetail
*textstructure:* floatingText
*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme
character data

*Note* Labels specifically relating to usage should be tagged with the special-purpose <usg> element rather than with the generic<lbl> element.

*Example*

```
<entry>
  <form  type="abbrev">
    <orth>MTBF</orth>
  </form>
  <form  type="full">
    <lbl>abbrev. for</lbl>
    <orth>mean time between failures</orth>
  </form>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element lbl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.paraContent}
```

---

## <leaf> (leaf) encodes the leaves (terminal nodes) of a tree. [*20.2. Trees*]

*Module* nets — Graphs, Networks, and Trees

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

* *@ana*

– att.global.change
* *@change*

– att.global.facs
* *@facs*

– att.global.linking
* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@value` provides a pointer to a feature structure or other analytic element.
*Status* Optional
*Datatype* teidata.pointer

`@parent` provides the identifier of parent of a leaf.
*Status* Optional
*Datatype* teidata.pointer

`@follow` provides an identifier of an element which this leaf follows.
*Status* Optional
*Datatype* teidata.pointer
*Note* If the tree is unordered or partially ordered, this attribute has the property of fixing the relative order of the leaf and the element which is the value of the attribute.

*Contained by* *tei:* tree
*May contain*
*core:* label
*Note* The in degree of a leaf is always 1, its out degree always 0.
*Example*

```
<leaf xml:id="peri1" parent="#n1">
  <label>periscope</label>
</leaf>
```

*Content model*

```
<content>
 <elementRef key="label" minOccurs="0"/>
</content>
```

*Schema Declaration*

```
element leaf
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute value { text }?,
   attribute parent { text }?,
   attribute follow { text }?,
   label?
}
```

---

**<lem>** (lemma) contains the lemma, or base text, of a textual variation. [*13.1. The Apparatus Entry, Readings, and Witnesses*]

*Module* textcrit — Critical Apparatus

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.textCritical
  - *@type*
  - *@cause*
  - *@varSeq*
  - *@require*

  - att.typed
    * ~~type~~
    * @subtype

  - att.written
    * *@hand*
- att.witnessed
  - *@wit*

*Contained by:* *tei:*lit: app rdgGrp

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList epilogue move performance prologue set sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app lacunaEnd lacunaStart listApp listWit wit witDetail witEnd witStart

*textstructure:* argument byline div docAuthor docDate docEdition docImprint docTitle epigraph floatingText imprimatur titlePage titlePart

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The term *lemma* is used in text criticism to describe the reading given in the main text, which may be used as a heading in the apparatus itself. This usage connects it to mathematics (where a lemma is a proven proposition used as a step in a proof, a ”given”) and natural-language processing (where a lemma is the dictionary headword associated with an inflected form in the running text).

*Example*

```
<app>
  <lem  wit="#El #Hg">Experience</lem>
  <rdg  wit="#La"  type="substantive">Experiment</rdg>
  <rdg  wit="#Ra2"  type="substantive">Eryment</rdg>
</app>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.divLike"/>
  <classRef key="model.divPart"/>
  <elementRef key="titlePage"/>
  <elementRef key="argument"/>
  <elementRef key="byline"/>
  <elementRef key="docAuthor"/>
  <elementRef key="docDate"/>
  <elementRef key="docEdition"/>
  <elementRef key="docImprint"/>
  <elementRef key="docTitle"/>
  <elementRef key="epigraph"/>
  <elementRef key="imprimatur"/>
  <elementRef key="titlePart"/>
  <elementRef key="epilogue"/>
  <elementRef key="performance"/>
  <elementRef key="prologue"/>
  <elementRef key="set"/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.global"/>
  <classRef key="model.rdgPart"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element lem
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.textCritical.attributes,
   att.written.attributes,
   att.witnessed.attributes,
   (
      text
    | model.divLike    | model.divPart    | titlePage    | argument    | byline    | docAu
}
```

---

**<lg>** (line group) contains one or more verse lines functioning as a formal unit, e.g. a stanza, refrain, verse paragraph, etc. [*3.13.1. Core Tags for Verse 3.13. Passages of Verse or Drama 7.2.5. Speech Contents*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*

    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.declaring
  - *@decls*
- att.divLike
  - *@org*
  - *@sample*
  - att.fragmentable
    * *@part*
  - att.metrical
    * *@met*
    * *@real*
    * *@rhyme*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.divPart model.paraPart

*Contained by*

*cmc:* post

*core:* add corr del emph head hi item lg note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def gen gram hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figure

*header:* change handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*tagdocs:* specGrp

*textcrit:* lem rdg

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* add cb corr del desc ellipsis gap gb head index l label lb lg meeting milestone note noteGrp orig pb reg sic stage unclear

*drama:* camera caption move sound tech view

*figures:* figure notatedMusic

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*textstructure:* argument byline closer dateline docAuthor docDate epigraph opener postscript salute signed trailer

*transcr:* addSpan damage damageSpan delSpan fw handShift listTranspose metamark mod redo restore retrace secl space substJoin supplied surplus undo

*Note* contains verse lines or nested line groups only, possibly prefixed by a heading.

*Example*

```
<lg  type="free">
  <l>Let me be my own fool</l>
  <l>of my own making, the sum of it</l>
</lg>
<lg  type="free">
  <l>is equivocal.</l>
  <l>One says of the drunken farmer:</l>
</lg>
<lg  type="free">
  <l>leave him lay off it. And this is</l>
  <l>the explanation.</l>
</lg>
```

*Schematron* <sch:rule context="tei:lg"> <sch:assert test="count(descendant::tei:lg|descendant::tei:l|descendant::tei:gap) > 0">An lg element must contain at least one child l, lg, or gap element.</sch:assert> </sch:rule>

*Schematron* <sch:rule context="tei:lg"> <sch:report test="ancestor::tei:l[not(.//tei:note//tei:lg[. = current()])]">Abstract model violation: Lines may not contain line groups.</sch:report> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <alternate>
   <classRef key="model.lLike"/>
   <classRef key="model.stageLike"/>
   <classRef key="model.labelLike"/>
   <classRef key="model.pPart.transcriptional"/>
   <elementRef key="lg"/>
  </alternate>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.lLike"/>
   <classRef key="model.stageLike"/>
   <classRef key="model.labelLike"/>
```

```
   <classRef key="model.pPart.transcriptional"/>
   <classRef key="model.global"/>
   <elementRef key="lg"/>
  </alternate>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divBottom"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element lg
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declaring.attributes,
   att.divLike.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   (
      ( model.divTop | model.global )*,
      (
         model.lLike       | model.stageLike       | model.labelLike       | model.pPart.t
      (
         model.lLike       | model.stageLike       | model.labelLike       | model.pPart.t
      ( ( model.divBottom, model.global* )* )
   )
}
```

---

## <licence>

contains information about a licence or other legal agreement applicable to the text. [*2.2.4. Publication, Distribution, Licensing, etc.*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*

    * *@notAfter*
    * *@from*
    * *@to*
  - att.pointing
    - *@targetLang*
    - *@target*
    - *@evaluate*

*Member of* model.availabilityPart

*Contained by*

*header:* availability

*linking:* annotation

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* A <licence> element should be supplied for each licence agreement applicable to the text in question. The *target* attribute may be used to reference a full version of the licence. The *when*, *notBefore*, *notAfter*, *from* or *to* attributes may be used in combination to indicate the date or dates of applicability of the licence.

*Example*

```
<licence target="http://www.nzetc.org/tm/scholarly/tei-NZETC-Help.html#licensing">
Licence: Creative Commons Attribution-Share Alike 3.0 New Zealand Licence
</licence>
```

*Example*

```
<availability>
  <licence target="http://creativecommons.org/licenses/by/3.0/"
    notBefore="2013-01-01">
    <p>The Creative Commons Attribution 3.0 Unported (CC BY 3.0) Licence
        applies to this document.</p>
    <p>The licence was added on January 1, 2013.</p>
  </licence>
</availability>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element licence
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.pointing.attributes,
   macro.specialPara}
```

---

**<line>** contains the transcription of a topographic line in the source document. [*12.2.2. Embedded Transcription*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.coordinated
  - *@start*
  - *@ulx*
  - *@uly*
  - *@lrx*
  - *@lry*
  - *@points*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*

*Member of* model.linePart
*Contained by*
*transcr:* line surface zone
*May contain*
*analysis:* c interp interpGrp pc span spanGrp w
*certainty:* certainty precision respons
*core:* add cb choice del ellipsis gap gb hi index lb milestone note noteGrp pb q unclear
*figures:* figure notatedMusic
*gaiji:* g

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*transcr:* addSpan damage damageSpan delSpan fw handShift line listTranspose metamark mod path redo restore retrace space substJoin undo zone

character data

*Note* This element should be used only to mark up writing which is topographically organized as a series of lines, horizontal or vertical. It should not be used to mark lines of verse (for which use <l>) nor to mark line beginnings within text which has been encoded using structural elements such as <p> (for which use <lb>).

*Example* This example shows topographical lines as a means of preserving the visual appearance of a poem:

```
<surface>
  <zone>
    <line>Poem</line>
    <line>As in Visions of – at</line>
    <line>night –</line>
    <line>All sorts of fancies running through</line>
    <line>the head</line>
  </zone>
</surface>
```

*Example*

```
<surface>
  <zone>
    <line>Hope you enjoyed</line>
    <line>Wales, as they
        said</line>
    <line>to Mrs FitzHerbert</line>
    <line>Mama</line>
  </zone>
  <zone>
    <line>Printed in England</line>
  </zone>
</surface>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.global"/>
  <classRef key="model.gLike"/>
  <classRef key="model.linePart"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element line
{
   att.global.attributes,
   att.global.analytic.attributes,
```

```
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.coordinated.attributes,
   att.typed.attributes,
   att.written.attributes,
   ( text | model.global | model.gLike | model.linePart )*
}
```

---

**<link>** (link) defines an association or hypertextual link among elements or passages, of some type not more precisely specifiable by other elements. [*17.1. Links*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

    * *@source*
  - att.cmc
    - *@generatedBy*
  - att.pointing
    - *@targetLang*
    - *@target*
    - *@evaluate*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.global.meta

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab linkGrp seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education event eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality object objectName occupation offset org orgName persName persPronouns person personGrp persona place placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Note* This element should only be used to encode associations not otherwise provided for by more specific elements.

The location of this element within a document has no significance, unless it is included within a <linkGrp>, in which case it may inherit the value of the *type* attribute from the value given on the <linkGrp>.

*Example*

```
<s  n="1">The state Supreme Court has refused to release <rs  xml:id="R1">
   <rs  xml:id="R2">Rahway State Prison</rs> inmate</rs>
  <rs  xml:id="R3">James Scott</rs> on bail.</s>
<s  n="2">
  <rs  xml:id="R4">The fighter</rs> is serving 30-40 years
for a 1975 armed robbery conviction in <rs  xml:id="R5">the penitentiary</rs>.
</s>
<!-- ... -->
<linkGrp  type="periphrasis">
  <link  target="#R1 #R3 #R4"/>
  <link  target="#R2 #R5"/>
</linkGrp>
```

*Schematron* <sch:rule context="tei:link"> <sch:assert test="contains(normalize-space(@target),' ')">You must supply at least two values for @target or on <sch:name/> </sch:assert> </sch:rule>

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element link
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.pointing.attributes,
   att.typed.attributes,
   empty
}
```

---

## <linkGrp>

(link group) defines a collection of associations or hypertextual links. [*17.1. Links*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change

    * *@change*
  – att.global.facs
    * *@facs*
  – att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  – att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  – att.global.responsibility
    * *@cert*
    * *@resp*
  – att.global.source
    * *@source*
- att.cmc
  – *@generatedBy*
- att.pointing.group
  – *@domains*
  – *@targFunc*

  – att.pointing
    * *@targetLang*
    * *@target*
    * *@evaluate*
  – att.typed
    * *@type*
    * *@subtype*

*Member of* model.global.meta

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education event eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality object objectName occupation offset org orgName persName persPronouns person personGrp persona place placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*core:* desc ptr

*linking:* link

*Note* May contain one or more <link> or <ptr> elements.

A web or link group is an administrative convenience, which should be used to collect a set of links together for any purpose, not simply to supply a default value for the *type* attribute.

*Example*

```
<linkGrp  type="translation">
  <link  target="#CCS1 #SW1"/>
  <link  target="#CCS2 #SW2"/>
  <link  target="#CCS #SW"/>
</linkGrp>
<div  type="volume"  xml:id="CCS"
  xml:lang="fr">
  <p>
    <s  xml:id="CCS1">Longtemps, je me suis couché de bonne heure.</s>
    <s  xml:id="CCS2">Parfois, à peine ma bougie éteinte, mes yeux se fermaient
si vite que je n'avais pas le temps de me dire : "Je m'endors."</s>
  </p>
<!-- ... -->
</div>
<div  type="volume"  xml:id="SW"  xml:lang="en">
  <p>
    <s  xml:id="SW1">For a long time I used to go to bed early.</s>
    <s  xml:id="SW2">Sometimes, when I had put out my candle, my eyes would
```

```
    close so quickly that I had not even time to say "I'm going to sleep."</s>
     </p>
    <!-- ... -->
    </div>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.descLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="link"/>
   <elementRef key="ptr"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element linkGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.pointing.group.attributes,
   att.pointing.attributes,
   att.typed.attributes,
   ( model.descLike*, ( link | ptr )+ )
}
```

---

## <list> (list) contains any sequence of items organized as a list. [*3.8. Lists*]

*Module* core — Elements Available in All TEI Documents

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change

  * *@change*
- att.global.facs
  * *@facs*
- att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

- att.cmc
  - *@generatedBy*
- att.sortable
  - *@sortKey*
- att.typed
  - ~~type~~
  - @subtype

`@type` (type) describes the nature of the items in the list.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **gloss** (gloss) each list item glosses some term or concept, which is given by a <label> element preceding the list item.

**index** (index) each list item is an entry in an index such as the alphabetical topical index at the back of a print volume.

**instructions** (instructions) each list item is a step in a sequence of instructions, as in a recipe.

**litany** (litany) each list item is one of a sequence of petitions, supplications or invocations, typically in a religious ritual.

**syllogism** (syllogism) each list item is part of an argument consisting of two or more propositions and a final conclusion derived from them.

*Note* Previous versions of these Guidelines recommended the use of *type* on <list> to encode the rendering or appearance of a list (whether it was bulleted, numbered, etc.). The current recommendation is to use the *rend* or *style* attributes for these

aspects of a list, while using *type* for the more appropriate task of characterizing the nature of the content of a list.

The formal syntax of the element declarations allows <label> tags to be omitted from lists tagged <list type="gloss">; this is however a semantic error.

*Member of* model.listLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote keywords licence rendition revisionDesc scriptNote sourceDesc tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* cb desc ellipsis gap gb head headItem headLabel index item label lb meeting milestone note noteGrp pb

*figures:* figure notatedMusic

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*textstructure:* argument byline closer dateline docAuthor docDate epigraph opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* May contain an optional heading followed by a series of items, or a series of label and item pairs, the latter being optionally preceded by one or two specialized headings.

*Example*

```
<list  rend="numbered">
  <item>a butcher</item>
  <item>a baker</item>
  <item>a candlestick maker, with
  <list  rend="bulleted">
     <item>rings on his fingers</item>
     <item>bells on his toes</item>
   </list>
```

```
  </item>
</list>
```

*Example*

```
<list  type="syllogism"  rend="bulleted">
  <item>All Cretans are liars.</item>
  <item>Epimenides is a Cretan.</item>
  <item>ERGO Epimenides is a liar.</item>
</list>
```

*Example*

```
<list  type="litany"  rend="simple">
  <item>God save us from drought.</item>
  <item>God save us from pestilence.</item>
  <item>God save us from wickedness in high places.</item>
  <item>Praise be to God.</item>
</list>
```

*Example* The following example treats the short numbered clauses of Anglo-Saxon legal codes as lists of items. The text is from an ordinance of King Athelstan (924–939):

```
<div1  type="section">
  <head>Athelstan's Ordinance</head>
  <list  rend="numbered">
    <item  n="1">Concerning thieves. First, that no thief is to be spared who is
caught with
       the stolen goods, [if he is] over twelve years and [if the value of the
goods is] over
       eightpence.
   <list  rend="numbered">
      <item  n="1.1">And if anyone does spare one, he is to pay for the thief
with his
              wergild – and the thief is to be no nearer a settlement on that
account – or to
              clear himself by an oath of that amount.</item>
      <item  n="1.2">If, however, he [the thief] wishes to defend himself or to
escape, he is
              not to be spared [whether younger or older than twelve].</item>
      <item  n="1.3">If a thief is put into prison, he is to be in prison 40
days, and he may
              then be redeemed with 120 shillings; and the kindred are to stand
surety for him
              that he will desist for ever.</item>
      <item  n="1.4">And if he steals after that, they are to pay for him with
his wergild,
              or to bring him back there.</item>
      <item  n="1.5">And if he steals after that, they are to pay for him with
his wergild,
              whether to the king or to him to whom it rightly belongs; and
everyone of those who
              supported him is to pay 120 shillings to the king as a
fine.</item>
     </list>
   </item>
   <item  n="2">Concerning lordless men. And we pronounced about these lordless
```

```
men, from whom
        no justice can be obtained, that one should order their kindred to
fetch back such a
        person to justice and to find him a lord in public meeting.
   <list  rend="numbered">
      <item  n="2.1">And if they then will not, or cannot, produce him on that
appointed day,
               he is then to be a fugitive afterwards, and he who encounters him
is to strike him
               down as a thief.</item>
      <item  n="2.2">And he who harbours him after that, is to pay for him with
his wergild
               or to clear himself by an oath of that amount.</item>
    </list>
   </item>
   <item  n="3">Concerning the refusal of justice. The lord who refuses justice
and upholds
        his guilty man, so that the king is appealed to, is to repay the value
of the goods and
        120 shillings to the king; and he who appeals to the king before he
demands justice as
        often as he ought, is to pay the same fine as the other would have
done, if he had
        refused him justice.
   <list  rend="numbered">
      <item  n="3.1">And the lord who is an accessory to a theft by his slave,
and it becomes
               known about him, is to forfeit the slave and be liable to his
wergild on the first
               occasionp if he does it more often, he is to be liable to pay all
that he owns.</item>
      <item  n="3.2">And likewise any of the king's treasurers or of our
reeves, who has been
               an accessory of thieves who have committed theft, is to liable to
the same.</item>
    </list>
   </item>
   <item  n="4">Concerning treachery to a lord. And we have pronounced
concerning treachery to
        a lord, that he [who is accused] is to forfeit his life if he cannot
deny it or is
        afterwards convicted at the three-fold ordeal.</item>
  </list>
</div1>
```

Note that nested lists have been used so the tagging mirrors the structure indicated by the two-level numbering of the clauses. The clauses could have been treated as a one-level list with irregular numbering, if desired.

*Example*

```
<p>These decrees, most blessed Pope Hadrian, we propounded in the public
council ... and they
confirmed them in our hand in your stead with the sign of the Holy Cross, and
afterwards
inscribed with a careful pen on the paper of this page, affixing thus the sign
of the Holy
```

```
Cross.
<list  rend="simple">
   <item>I, Eanbald, by the grace of God archbishop of the holy church of
York, have
       subscribed to the pious and catholic validity of this document with the
sign of the Holy
       Cross.</item>
   <item>I, Ælfwold, king of the people across the Humber, consenting have
subscribed with
       the sign of the Holy Cross.</item>
   <item>I, Tilberht, prelate of the church of Hexham, rejoicing have
subscribed with the
       sign of the Holy Cross.</item>
   <item>I, Higbald, bishop of the church of Lindisfarne, obeying have
subscribed with the
       sign of the Holy Cross.</item>
   <item>I, Ethelbert, bishop of Candida Casa, suppliant, have subscribed with
thef sign of
       the Holy Cross.</item>
   <item>I, Ealdwulf, bishop of the church of Mayo, have subscribed with
devout will.</item>
   <item>I, Æthelwine, bishop, have subscribed through delegates.</item>
   <item>I, Sicga, patrician, have subscribed with serene mind with the sign
of the Holy
       Cross.</item>
 </list>
</p>
```

*Schematron* <sch:rule context="tei:list[@type='gloss']"> <sch:assert test="tei:label">The content of a "gloss" list should include a sequence of one or more pairs of a label element followed by an item element</sch:assert> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
   <elementRef key="desc" minOccurs="0"
    maxOccurs="unbounded"/>
  </alternate>
  <alternate>
   <sequence minOccurs="1"
    maxOccurs="unbounded">
    <elementRef key="item"/>
    <classRef key="model.global"
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
   <sequence>
    <elementRef key="headLabel"
     minOccurs="0"/>
    <elementRef key="headItem"
     minOccurs="0"/>
    <sequence minOccurs="1"
     maxOccurs="unbounded">
     <elementRef key="label"/>
     <classRef key="model.global"
```

```
      minOccurs="0" maxOccurs="unbounded"/>
     <elementRef key="item"/>
     <classRef key="model.global"
      minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
   </sequence>
  </alternate>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divBottom"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element list
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.sortable.attributes,
   att.typed.attribute.subtype,
   attribute type
   {
      "gloss" | "index" | "instructions" | "litany" | "syllogism"
   }?,
   (
      ( model.divTop | model.global | desc* )*,
      (
         ( ( item, model.global* )+ )
       | (
            headLabel?,
            headItem?,
            ( ( label, model.global*, item, model.global* )+ )
         )
      ),
      ( ( model.divBottom, model.global* )* )
   )
}
```

---

**<listAnnotation>** contains a list of annotations, typically encoded as <annotation>, <annotationBlock>, or <note>, possibly organized with nested <listAnnotation> elements. [*17.10. The standOff Container*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*
- att.global
  - *@xml:id*

  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declaring
  - *@decls*
- att.notated
  - *@notation*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.standOffPart
*Contained by*
*linking:* listAnnotation standOff
*May contain*
*core:* desc head label note
*linking:* annotation listAnnotation
*spoken:* annotationBlock

*Note* The enclosed annotations may use the general-purpose <note> element; or, for annotations pertaining to transcriptions of speech, the special-purpose <annotationBlock> element; or the <annotation> element, which is intended to map cleanly onto the Web Annotation Data Model.

*Example*

```
<standOff>
  <listAnnotation>
    <note  target="#RotAM.4.15"  place="margin"
      resp="#STC"  type="gloss"> The spell begins to
         break </note>
    <note  target="#RotAM.4.15"  place="bottom"
      resp="#JLL"> The turning point of the poem...
    </note>
  </listAnnotation>
</standOff>
<!-- ... -->
<lg  xml:id="RotAM.4.15"  rhyme="ABCB">
  <l>The self-same moment I could pray;</l>
  <l>And from my neck so free</l>
  <l>The albatross fell off, and sank</l>
  <l>Like lead into the sea.</l>
</lg>
```

*Example*

```
<listAnnotation>
  <annotationBlock  corresp="#u1">
    <spanGrp  type="pos">
      <span  from="#e14.w1"  to="#e14.w1">CONJ</span>
      <span  from="#e14.w2"  to="#e14.w2">RELPRO</span>
      <span  from="#e14.w3"  to="#e14.w3">V</span>
    </spanGrp>
    <spanGrp  type="prosody">
      <span  from="#T0"  to="#T1">faster</span>
    </spanGrp>
  </annotationBlock>
<!-- ... annotationBlocks for #u2 through #u27 here -->
</listAnnotation>
```

*Example*

```
<listAnnotation>
  <annotation  xml:id="bgann1"
    motivation="commenting"
    target="#match(bg-c1p1s1,'Gallia.*omnis')">
    <respStmt>
      <resp>creator</resp>
      <persName>Francis Kelsey</persName>
    </respStmt>
    <note>‘Gaul as a whole,’ contrasted with Gaul in the narrower sense, or
Celtic Gaul; Celtic Gaul also is often called Gallia.</note>
  </annotation>
  <annotation  xml:id="bgann2"
    motivation="commenting"
    target="#match(bg-c1p1s1,'Gallia.*divisa')">
    <respStmt>
      <resp>creator</resp>
```

```
      <persName>Rice Holmes</persName>
    </respStmt>
    <note>Gallia...divisa: Notice the order of the words. ‘Gaul, taken as a
whole, is divided’.</note>
  </annotation>
  <annotation  xml:id="bgann3"
    motivation="commenting"  target="#match(bg-c1p1s1,'Belgae')">
    <respStmt>
      <resp>creator</resp>
      <persName>Arthur Tappan Walker</persName>
    </respStmt>
    <note>Belgae -arum m., the Belgae or Belgians</note>
  </annotation>
  <annotation  xml:id="bgann4"
    motivation="commenting"  target="#match(bg-c1p1s1,'Aquitani')">
    <respStmt>
      <resp>creator</resp>
      <persName>Arthur Tappan Walker</persName>
    </respStmt>
    <note>Aquitani, -orum m.: the Aquitani, inhabiting southwestern Gaul</note>
  </annotation>
  <annotation  xml:id="bgann5"
    motivation="commenting"  target="#match(bg-c1p1s1,'Celtae')">
    <respStmt>
      <resp>creator</resp>
      <persName>Arthur Tappan Walker</persName>
    </respStmt>
    <note>Celtae, -arum m: the Celtae or Celts</note>
  </annotation>
  <annotation  xml:id="bgann6"
    motivation="commenting"
    target="#match(bg-c1p1s2,'Gallos(.|\n)*dividit')">
    <respStmt>
      <resp>creator</resp>
      <persName>William Francis Allen</persName>
      <persName>Joseph Henry Allen</persName>
      <persName>Harry Pratt Judson</persName>
    </respStmt>
    <note>the verb is singular, because the two rivers make one boundary; as we
should say,
        ‘is divided by the line of the Seine and Marne.’</note>
  </annotation>
</listAnnotation>
<!-- Elsewhere in the document -->
<text>
  <body>
    <div  type="edition">
      <div  type="textpart"  subtype="chapter"
        n="1"  xml:id="bg-c1">
        <p  n="1"  xml:id="bg-c1p1">
          <seg  n="1"  xml:id="bg-c1p1s1">Gallia est omnis divisa in partes tres,
quarum unam incolunt Belgae, aliam
                  Aquitani, tertiam qui ipsorum lingua Celtae, nostra Galli
appellantur.</seg>
          <seg  n="2"  xml:id="bg-c1p1s2">Hi omnes lingua, institutis, legibus
inter se differunt. Gallos ab Aquitanis
                  Garumna flumen, a Belgis Matrona et Sequana dividit.</seg>
```

```
    <!-- ... -->
          </p>
        </div>
      </div>
    </body>
  </text>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <classRef key="model.labelLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.annotationLike"/>
   <elementRef key="listAnnotation"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element listAnnotation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.notated.attributes,
   att.typed.attributes,
   (
      model.headLike*,
      model.labelLike*,
      ( model.annotationLike | listAnnotation )+
   )
}
```

---

**<listApp>** (list of apparatus entries) contains a list of apparatus entries. [*13.2. Linking the Apparatus to the Text*]

*Module* textcrit — Critical Apparatus

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.declarable
  - *@default*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.listLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote licence rendition scriptNote sourceDesc tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*textcrit:* lem listApp rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc head

*textcrit:* app listApp

*Note* <listApp> elements would normally be located in the <back> part of a document, but they may appear elsewhere.

*Example* In the following example from the exegetical Yasna, the base text is encoded in the <body> of the document, and two separate <listApp> elements are used in the <back>, containing variant readings written in different scripts.

```
<body>
  <div>
    <lg type="stanza" xml:id="Y-36.01"
      xml:lang="pal-Avst" rend="italic">
      <l xml:id="Y-36.01_L-1">
        <w xml:id="Y-36.01_L1_W-01">ahiiā</w>
        <w xml:id="Y-36.01_L1_W-02">θβā</w>
        <w xml:id="Y-36.01_L1_W-03">āθrō</w>
        <w xml:id="Y-36.01_L1_W-04">vərəzānā</w>
        <w xml:id="Y-36.01_L1_W-05">paouruiiē</w>
        <w xml:id="Y-36.01_L1_W-06">pairijasāmaiδē</w>
        <w xml:id="Y-36.01_L1_W-07">mazdā</w>
        <w xml:id="Y-36.01_L1_W-08">ahurā</w>
      </l>
<!-- ... -->
    </lg>
  </div>
</body>
<!-- ... -->
<back>
  <div>
    <listApp xml:id="CA_Y-36"
      xml:lang="pal-Avst">
      <head>Variants from witnesses in Avestan script</head>
      <app from="#Y-36.01_L1_W-01">
        <rdg wit="#Pt4 #F2 #J2 #M1">ahiiā</rdg>
      </app>
      <app from="#Y-36.01_L1_W-02">
        <rdg wit="#Pt4 #F2 #J2 #M1">θβā</rdg>
      </app>
      <app from="#Y-36.01_L1_W-03">
        <rdg wit="#Pt4 #J2 #M1">āθrō</rdg>
        <rdg wit="#F2">āθrōi</rdg>
      </app>
<!-- ... -->
```

```
      </listApp>
      <listApp  xml:id="CA_PY-36"
        xml:lang="pal-Phlv">
        <head>Variants from witnesses written in Pahlavi script</head>
        <app  from="#PY-36.01_L1_W-01">
          <rdg  wit="#Pt4 #F2 #J2 #M1">ʾytwn'</rdg>
        </app>
        <app  from="#PY-36.01_L1_W-02">
          <rdg  wit="#Pt4 #F2 #J2 #M1">ʾw'</rdg>
        </app>
        <app  from="#PY-36.01_L1_W-03">
          <rdg  wit="#Pt4 #F2 #J2 #M1">ḤNʾ</rdg>
        </app>
    <!-- ... -->
      </listApp>
    </div>
  </back>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:listApp"/> </sch:pattern>
*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="app"/>
   <elementRef key="listApp"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element listApp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   ( model.headLike*, desc*, ( app | listApp )+ )
}
```

**<listBibl>** (citation list) contains a list of bibliographic citations of any kind. [*3.12.1. Methods of Encoding Bibliographic References and Lists of References 2.2.7. The Source Description 16.3.2. Declarable Elements*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.declarable
  - *@default*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.biblLike model.frontPart

*Contained by*

*cmc:* post

*core:* add cit corr del desc emph head hi item l listBibl meeting note orig p q quote ref reg relatedItem said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote licence rendition scriptNote sourceDesc tagUsage taxonomy

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additional additions collation condition custEvent decoNote filiation foliation layout msItem msItemStruct musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* climate event location object occupation org person personGrp persona place population state terrain trait

*spoken:* broadcast scriptStmt u writing

*textcrit:* lem rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph front imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* bibl biblStruct cb desc gb head lb listBibl milestone pb

*header:* biblFull

*linking:* anchor

*msdescription:* msDesc

*namesdates:* listRelation relation

*transcr:* fw

*Example*

```
<listBibl>
  <head>Works consulted</head>
  <bibl>Blain, Clements and Grundy: Feminist Companion to
     Literature in English (Yale, 1990)
  </bibl>
  <biblStruct>
    <analytic>
      <title>The Interesting story of the Children in the Wood</title>
    </analytic>
    <monogr>
      <title>The Penny Histories</title>
      <author>Victor E Neuberg</author>
      <imprint>
        <publisher>OUP</publisher>
        <date>1968</date>
      </imprint>
    </monogr>
  </biblStruct>
</listBibl>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:listBibl"/> </sch:pattern>

*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.milestoneLike"
    minOccurs="1" maxOccurs="1"/>
   <elementRef key="relation" minOccurs="1"
    maxOccurs="1"/>
   <elementRef key="listRelation"
    minOccurs="1" maxOccurs="1"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.biblLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.milestoneLike"
     minOccurs="1" maxOccurs="1"/>
    <elementRef key="relation"
     minOccurs="1" maxOccurs="1"/>
    <elementRef key="listRelation"
     minOccurs="1" maxOccurs="1"/>
   </alternate>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element listBibl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declarable.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      model.headLike*,
      desc*,
      ( model.milestoneLike | relation | listRelation )*,
      (
         (
            model.biblLike+,
            ( model.milestoneLike | relation | listRelation )*
```

```
            )+
        )
    )
}
```

---

**<listChange>** groups a number of change descriptions associated with either the creation of a source text or the revision of an encoded text. [*2.6. The Revision Description 12.7. Identifying Changes and Revisions*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*

– *@subtype*

`@ordered` indicates whether the ordering of its child <change> elements is to be considered significant or not.
*Status* Optional
*Datatype* teidata.truthValue
*Default* true

*Member of* model.standOffPart
*Contained by*
*header:* creation listChange revisionDesc
*linking:* standOff
*May contain*
*core:* desc
*header:* change listChange
*Note* When this element appears within the <creation> element it documents the set of revision campaigns or stages identified during the evolution of the original text. When it appears within the <revisionDesc> element, it documents only changes made during the evolution of the encoded representation of that text.
*Example*

```
<revisionDesc>
  <listChange>
    <change  when="1991-11-11"  who="#LB"> deleted chapter 10 </change>
    <change  when="1991-11-02"  who="#MSM"> completed first draft </change>
  </listChange>
</revisionDesc>
```

*Example*

```
<profileDesc>
  <creation>
    <listChange  ordered="true">
      <change  xml:id="CHG-1">First stage, written in ink by a writer</change>
      <change  xml:id="CHG-2">Second stage, written in Goethe's hand using
pencil</change>
      <change  xml:id="CHG-3">Fixation of the revised passages and further
revisions by
            Goethe using ink</change>
      <change  xml:id="CHG-4">Addition of another stanza in a different hand,
            probably at a later stage</change>
    </listChange>
  </creation>
</profileDesc>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="listChange"/>
   <elementRef key="change"/>
  </alternate>
 </sequence>
```

```
</content>
```

*Schema Declaration*

```
element listChange
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   attribute ordered { text }?,
   ( desc*, ( listChange | change )+ )
}
```

---

## <listEvent>

<listEvent> (list of events) contains a list of descriptions, each of which provides information about an identifiable event. [*14.3.1. Basic Principles*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

    * *@rend*
    * *@style*
    * *@rendition*
  – att.global.responsibility
    * *@cert*
    * *@resp*
  – att.global.source
    * *@source*
- att.cmc
  – *@generatedBy*
- att.declarable
  – *@default*
- att.sortable
  – *@sortKey*
- att.typed
  – *@type*
  – *@subtype*

*Member of* model.eventLike model.listLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote licence rendition scriptNote sourceDesc tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* event listEvent occupation org person personGrp persona place

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc head

*namesdates:* event listEvent listRelation relation

*Example*

```
<listEvent>
  <head>Battles of the American Civil War: Kentucky</head>
  <event  xml:id="event01"  when="1861-09-19">
```

```
      <label>Barbourville</label>
      <desc>The Battle of Barbourville was one of the early engagements of
          the American Civil War. It occurred September 19, 1861, in Knox
          County, Kentucky during the campaign known as the Kentucky Confederate
          Offensive. The battle is considered the first Confederate victory in
          the commonwealth, and threw a scare into Federal commanders, who
          rushed troops to central Kentucky in an effort to repel the invasion,
          which was finally thwarted at the <ref  target="#event02">Battle of
             Camp Wildcat</ref> in October.</desc>
    </event>
    <event  xml:id="event02"  when="1861-10-21">
      <label>Camp Wild Cat</label>
      <desc>The Battle of Camp Wildcat (also known as Wildcat Mountain and Camp
          Wild Cat) was one of the early engagements of the American Civil
          War. It occurred October 21, 1861, in northern Laurel County, Kentucky
          during the campaign known as the Kentucky Confederate Offensive. The
          battle is considered one of the very first Union victories, and marked
          the first engagement of troops in the commonwealth of Kentucky.</desc>
    </event>
    <event  xml:id="event03"  from="1864-06-11"
      to="1864-06-12">
      <label>Cynthiana</label>
      <desc>The Battle of Cynthiana (or Kellar’s Bridge) was an engagement
          during the American Civil War that was fought on June 11 and 12, 1864,
          in Harrison County, Kentucky, near the town of Cynthiana. A part of
          Confederate Brigadier General John Hunt Morgan's 1864 Raid into
          Kentucky, the battle resulted in a victory by Union forces over the
          raiders and saved the town from capture.</desc>
    </event>
  </listEvent>
```

*Schematron* <sch:pattern is-a=”declarable”> <sch:param name=”tde” value=”tei:listEvent”/> </sch:pattern>
*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="relation" minOccurs="1"
    maxOccurs="1"/>
   <elementRef key="listRelation"
    minOccurs="1" maxOccurs="1"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.eventLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="relation"
     minOccurs="1" maxOccurs="1"/>
    <elementRef key="listRelation"
     minOccurs="1" maxOccurs="1"/>
   </alternate>
```

```
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element listEvent
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declarable.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      model.headLike*,
      desc*,
      ( relation | listRelation )*,
      ( ( model.eventLike+, ( relation | listRelation )* )+ )
   )
}
```

---

**<listForest>** provides for lists of forests. [*20.3. Another Tree Notation*]

*Module* nets — Graphs, Networks, and Trees

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*

  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

- att.typed
  - ~~type~~
  - @subtype

`@type` identifies the type of the forest group.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated

*Member of* model.divPart model.standOffPart
*Contained by*
*core:* item note q quote said stage textLang
*drama:* castList epilogue performance prologue set view
*figures:* cell figure
*header:* change handNote licence scriptNote
*linking:* standOff
*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote
*namesdates:* occupation
*tagdocs:* specGrp
*textcrit:* lem rdg
*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript
*transcr:* metamark
*May contain*
*core:* desc
*nets:* forest
*Note* One or more forests representing the same object.
*Example*

```
<listForest>
  <forest type="derivation-syntactic">
    <tree ord="true">
      <root children="#fgex1 #fgex2"/>
<!-- ... -->
    </tree>
```

```
<!-- ... -->
  </forest>
  <forest type="derivation-prosodic">
    <tree ord="true">
      <root children="#fgex3 #fgex4"/>
<!-- ... -->
    </tree>
<!-- ... -->
  </forest>
</listForest>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="forest" minOccurs="1"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element listForest
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   ( desc*, forest+ )
}
```

---

## <listNym>

(list of canonical names) contains a list of nyms, that is, standardized names for any thing. [*14.3.7. Names and Nyms*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.declarable
  - *@default*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.listLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote licence rendition scriptNote sourceDesc tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* listNym occupation
*spoken:* u writing
*textcrit:* lem rdg witness
*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer
*transcr:* damage metamark mod restore retrace secl supplied surplus
*verse:* rhyme
*May contain*
*core:* desc head
*namesdates:* listNym listRelation nym relation
*Note* The type attribute may be used to distinguish lists of names of a particular type if convenient.
*Example*

```
<listNym type="floral">
  <nym xml:id="ROSE">
    <form>Rose</form>
  </nym>
  <nym xml:id="DAISY">
    <form>Daisy</form>
    <etym>Contraction of <mentioned>day's eye</mentioned>
    </etym>
  </nym>
  <nym xml:id="HTHR">
    <form>Heather</form>
  </nym>
</listNym>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:listNym"/> </sch:pattern>
*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="relation" minOccurs="1"
    maxOccurs="1"/>
   <elementRef key="listRelation"
    minOccurs="1" maxOccurs="1"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <alternate minOccurs="1"
    maxOccurs="unbounded">
    <elementRef key="nym" minOccurs="1"
     maxOccurs="1"/>
    <elementRef key="listNym" minOccurs="1"
     maxOccurs="1"/>
   </alternate>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="relation"
     minOccurs="1" maxOccurs="1"/>
```

```
    <elementRef key="listRelation"
     minOccurs="1" maxOccurs="1"/>
   </alternate>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element listNym
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declarable.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      model.headLike*,
      desc*,
      ( relation | listRelation )*,
      ( ( ( nym | listNym )+, ( relation | listRelation )* )+ )
   )
}
```

---

## <listObject>

(list of objects) contains a list of descriptions, each of which provides information about an identifiable physical object. [*14.3.6. Objects*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.declarable
  - *@default*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.listLike model.objectLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote licence rendition scriptNote sourceDesc tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* event listObject occupation

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc head

*namesdates:* listObject listRelation object relation

*Note* The *type* attribute may be used to distinguish different types of objects.

*Example*

```
<listObject>
  <object  xml:id="AlfredJewel">
    <objectIdentifier>
      <country>United Kingdom</country>
      <region>Oxfordshire</region>
      <settlement>Oxford</settlement>
      <institution>University of Oxford</institution>
      <repository>Ashmolean Museum</repository>
      <collection>English Treasures</collection>
      <idno  type="ashmolean">AN1836p.135.371</idno>
      <idno  type="wikipedia">https://en.wikipedia.org/wiki/Alfred_Jewel</idno>
      <objectName>Alfred Jewel</objectName>
    </objectIdentifier>
    <physDesc>
      <p> The Alfred Jewel is about 6.4 cm in length and is made of combination
of filigreed <material>gold</material>
            surrounding a polished teardrop shaped piece of transparent
<material>quartz</material>. Underneath the rock crystal
            is a cloisonné enamel image of a man with ecclesiastical symbols.
The sides of the jewel holding the crystal in
            place contain an openwork inscription saying "AELFRED MEC HEHT
GEWYRCAN", meaning 'Alfred ordered me made'. </p>
    </physDesc>
    <history>
      <origin>It is generally accepted that the Alfred Jewel dates from the
<origDate>late 9th Century</origDate> and was
            most likely made in <origPlace>England</origPlace>. </origin>
      <provenance  when="1693">The jewel was discovered in 1693 at Petherton
Park, North Petherton in the English county of
            Somerset, on land owned by Sir Thomas Wroth. North Petherton is
about 8 miles away from Athelney, where King Alfred
            founded a monastery. </provenance>
      <provenance  when="1698">A description of the Alfred Jewel was first
published in 1698, in the Philosophical
            Transactions of the Royal Society.</provenance>
      <acquisition> It was bequeathed to Oxford University by Colonel Nathaniel
Palmer (c. 1661-1718) and today is in the
            Ashmolean Museum in Oxford. </acquisition>
    </history>
  </object>
</listObject>
```

*Schematron* <sch:pattern is-a=”declarable”> <sch:param name=”tde” value=”tei:listObject”/> </sch:pattern>

*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="desc" minOccurs="0"
```

```
   maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="relation" minOccurs="1"
    maxOccurs="1"/>
   <elementRef key="listRelation"
    minOccurs="1" maxOccurs="1"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.objectLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="relation"
     minOccurs="1" maxOccurs="1"/>
    <elementRef key="listRelation"
     minOccurs="1" maxOccurs="1"/>
   </alternate>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element listObject
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declarable.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      model.headLike*,
      desc*,
      ( relation | listRelation )*,
      ( ( model.objectLike+, ( relation | listRelation )* )+ )
   )
}
```

---

## <listOrg>

<listOrg> (list of organizations) contains a list of elements, each of which provides information about an identifiable organization. [*14.2.2. Organizational Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.declarable
  - *@default*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.listLike model.orgPart

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*corpus:* particDesc

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote licence rendition scriptNote sourceDesc tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* listOrg occupation org

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc head

*namesdates:* listOrg listRelation org relation

*Note* The type attribute may be used to distinguish lists of organizations of a particular type if convenient.

*Example*

```
<listOrg>
  <head>Libyans</head>
  <org>
    <orgName>Adyrmachidae</orgName>
    <desc>These people have, in most points, the same customs as the Egyptians,
but
        use the costume of the Libyans. Their women wear on each leg a ring
made of
        bronze [...]</desc>
  </org>
  <org>
    <orgName>Nasamonians</orgName>
    <desc>In summer they leave their flocks and herds upon the sea-shore, and
go up
        the country to a place called Augila, where they gather the dates from
the
        palms [...]</desc>
  </org>
  <org>
    <orgName>Garamantians</orgName>
    <desc>[...] avoid all society or intercourse with their fellow-men, have no
        weapon of war, and do not know how to defend themselves. [...]</desc>
<!-- ... -->
  </org>
</listOrg>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:listOrg"/> </sch:pattern>

*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
```

```
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="relation" minOccurs="1"
    maxOccurs="1"/>
   <elementRef key="listRelation"
    minOccurs="1" maxOccurs="1"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <alternate minOccurs="1"
    maxOccurs="unbounded">
    <elementRef key="org" minOccurs="1"
     maxOccurs="1"/>
    <elementRef key="listOrg" minOccurs="1"
     maxOccurs="1"/>
   </alternate>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="relation"
     minOccurs="1" maxOccurs="1"/>
    <elementRef key="listRelation"
     minOccurs="1" maxOccurs="1"/>
   </alternate>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element listOrg
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declarable.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      model.headLike*,
      desc*,
      ( relation | listRelation )*,
      ( ( ( org | listOrg )+, ( relation | listRelation )* )+ )
   )
}
```

**<listPerson>** (list of persons) contains a list of descriptions, each of which provides information about an identifiable person or a group of people, for example the participants in a language interaction, or the people referred to in a historical source. [*14.3.2. The Person Element 16.2. Contextual Information 2.4. The Profile Description 16.3.2. Declarable Elements*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.declarable
  - *@default*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*

– *@subtype*

*Member of* model.listLike model.orgPart

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*corpus:* particDesc

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote licence rendition scriptNote sourceDesc tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* event listPerson occupation org

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc head

*namesdates:* listPerson listRelation org person personGrp relation

*Note* The *type* attribute may be used to distinguish lists of people of a particular type if convenient.

*Example*

```
<listPerson type="respondents">
  <personGrp xml:id="PXXX"/>
  <person xml:id="P1234" sex="2" age="mid"/>
  <person xml:id="P4332" sex="1" age="mid"/>
  <listRelation>
    <relation type="personal" name="spouse"
     mutual="#P1234 #P4332"/>
  </listRelation>
</listPerson>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:listPerson"/> </sch:pattern>

*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="relation" minOccurs="1"
    maxOccurs="1"/>
```

```
   <elementRef key="listRelation"
    minOccurs="1" maxOccurs="1"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <alternate minOccurs="1"
    maxOccurs="unbounded">
    <classRef key="model.personLike"
     minOccurs="1" maxOccurs="1"/>
    <elementRef key="listPerson"
     minOccurs="1" maxOccurs="1"/>
   </alternate>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="relation"
     minOccurs="1" maxOccurs="1"/>
    <elementRef key="listRelation"
     minOccurs="1" maxOccurs="1"/>
   </alternate>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element listPerson
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declarable.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      model.headLike*,
      desc*,
      ( relation | listRelation )*,
      ( ( ( model.personLike | listPerson )+, ( relation | listRelation )* )+ )
   )
}
```

---

## <listPlace>

(list of places) contains a list of places, optionally followed by a list of relationships (other than containment) defined amongst them. [*2.2.7. The Source Description 14.3.4. Places*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.declarable
  - *@default*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.listLike model.orgPart

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*corpus:* settingDesc

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote licence rendition scriptNote sourceDesc tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* event listPlace occupation org place

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc head

*namesdates:* listPlace listRelation place relation

*Example*

```
<listPlace type="offshoreIslands">
  <place>
    <placeName>La roche qui pleure</placeName>
  </place>
  <place>
    <placeName>Ile aux cerfs</placeName>
  </place>
</listPlace>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:listPlace"/> </sch:pattern>

*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="relation" minOccurs="1"
    maxOccurs="1"/>
   <elementRef key="listRelation"
    minOccurs="1" maxOccurs="1"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <alternate minOccurs="1"
    maxOccurs="unbounded">
    <classRef key="model.placeLike"
     minOccurs="1" maxOccurs="1"/>
    <elementRef key="listPlace"
     minOccurs="1" maxOccurs="1"/>
```

```
   </alternate>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="relation"
     minOccurs="1" maxOccurs="1"/>
    <elementRef key="listRelation"
     minOccurs="1" maxOccurs="1"/>
   </alternate>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element listPlace
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declarable.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      model.headLike*,
      desc*,
      ( relation | listRelation )*,
      ( ( ( model.placeLike | listPlace )+, ( relation | listRelation )* )+ )
   )
}
```

---

## <listPrefixDef>

(list of prefix definitions) contains a list of definitions of prefixing schemes used in `teidata.pointer` values, showing how abbreviated URIs using each scheme may be expanded into full URIs. [*17.2.3. Using Abbreviated Pointers*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.encodingDescPart

*Contained by*

*header:* encodingDesc listPrefixDef

*May contain*

*core:* desc

*header:* listPrefixDef prefixDef

*Example* In this example, two private URI scheme prefixes are defined and patterns are provided for dereferencing them. Each prefix is also supplied with a human-readable explanation in a <p> element.

```
<listPrefixDef>
  <prefixDef  ident="psn"
    matchPattern="([A-Z]+)"
    replacementPattern="personography.xml#$1">
    <p> Private URIs using the <code>psn</code>
        prefix are pointers to <gi>person</gi>
        elements in the personography.xml file.
        For example, <code>psn:MDH</code>
        dereferences to <code>personography.xml#MDH</code>.
    </p>
  </prefixDef>
  <prefixDef  ident="bibl"
    matchPattern="([a-z]+[a-z0-9]*)"
    replacementPattern="http://www.example.com/getBibl.xql?id=$1">
    <p> Private URIs using the <code>bibl</code> prefix can be
        expanded to form URIs which retrieve the relevant
        bibliographical reference from www.example.com.
    </p>
  </prefixDef>
</listPrefixDef>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="prefixDef"/>
   <elementRef key="listPrefixDef"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element listPrefixDef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( desc*, ( prefixDef | listPrefixDef )+ )
}
```

---

## <listRef>

<listRef> (list of references) supplies a list of significant references in the current document or elsewhere. [*23.4.1. Description of Components 24.5.1. Making a Unified ODD*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.ptrLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine analytic author bibl biblScope biblStruct cit citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned monogr name note num orig p pubPlace publisher q quote rb ref reg relatedItem resp rs rt said series sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc notatedMusic

*header:* application authority catDesc change classCode correspContext creation distributor edition extent funder geoDecl handNote language licence principal publicationStmt rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education event eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person place placeName region residence roleName settlement sex socecStatus surname

*nets:* eLeaf eTree

*spoken:* transcriptionDesc u writing

*tagdocs:* classSpec dataSpec eg elementSpec listRef macroSpec moduleSpec schemaSpec specGrp valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*core:* desc ptr ref

*tagdocs:* listRef

*Note* When <listRef> is used in the context of `tagdocs`, further restrictons apply:

- It may only contain <ptr> or <ref> elements.
- Each <ptr> or <ref> child should bear a *target* attribute, the value of which should be a single URI pointer.

*Example*

```
<listRef>
  <ptr  target="#ddc12"/>
</listRef>
```

*Schematron* <sch:rule context="( tei:classSpec | tei:dataSpec | tei:elementSpec | tei:macroSpec | tei:moduleSpec | tei:schemaSpec | tei:specGrp )/tei:listRef"> <sch:report test="tei:listRef">In the context of tagset documentation, the listRef element must not self-nest.</sch:report> </sch:rule>

*Schematron* <sch:rule context="( tei:classSpec | tei:dataSpec | tei:elementSpec | tei:macroSpec | tei:moduleSpec | tei:schemaSpec | tei:specGrp )/tei:listRef/tei:ptr | ( tei:classSpec | tei:dataSpec | tei:elementSpec | tei:macroSpec | tei:moduleSpec | tei:schemaSpec | tei:specGrp )/tei:listRef/tei:ref"> <sch:assert test="@target and not( matches( @target,'\s') )">In the context of tagset documentation, each ptr or ref element inside a listRef must have a target attribute with only 1 pointer as its value.</sch:assert> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <classRef key="model.ptrLike"
   minOccurs="1" maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element listRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   ( desc*, model.ptrLike+ )
}
```

**<listRelation>** provides information about relationships identified amongst people, places, and organizations, either informally as prose or as formally expressed relation links. [*14.3.2.3. Personal Relationships*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.biblPart model.listLike

*Contained by*

*cmc:* post

*core:* add bibl corr del desc emph head hi item l listBibl meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote licence rendition scriptNote sourceDesc tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* event listEvent listNym listObject listOrg listPerson listPlace listRelation occupation

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc head p

*linking:* ab

*namesdates:* listRelation relation

*Note* May contain a prose description organized as paragraphs, or a sequence of <relation> elements.

*Example*

```
<listPerson>
  <person xml:id="pp1">
<!-- data about person pp1 -->
  </person>
  <person xml:id="pp2">
<!-- data about person pp1 -->
  </person>
<!-- more person (pp3, pp4) elements here -->
  <listRelation type="personal">
    <relation name="parent"
      active="#pp1 #pp2" passive="#pp3 #pp4"/>
    <relation name="spouse"
      mutual="#pp1 #pp2"/>
  </listRelation>
  <listRelation type="social">
    <relation name="employer" active="#pp1"
      passive="#pp3 #pp5 #pp6 #pp7"/>
  </listRelation>
</listPerson>
```

The persons with identifiers pp1 and pp2 are the parents of pp3 and pp4; they are also married to each other; pp1 is the employer of pp3, pp5, pp6, and pp7.

*Example*

```
<listPerson>
  <person  xml:id="en_pp1">
<!-- data about person en_pp1 -->
  </person>
  <person  xml:id="en_pp2">
<!-- data about person en_pp2 -->
  </person>
<!-- more person (en_pp3, en_pp4) elements here -->
</listPerson>
<listPlace>
  <place  xml:id="en_pl1">
<!-- data about place en_pl1 -->
  </place>
<!-- more place (en_pl2, en_pl3) elements here -->
</listPlace>
<listRelation>
  <relation  name="residence"
    active="#en_pp1 #en_pp2"  passive="#en_pl1"/>
</listRelation>
```

The persons with identifiers en_pp1 and en_pp2 live in en_pl1.

*Example*

```
<listRelation>
  <p>All speakers are members of the Ceruli family, born in Naples.</p>
</listRelation>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="1" maxOccurs="1">
   <classRef key="model.pLike"/>
   <alternate minOccurs="1"
    maxOccurs="unbounded">
    <elementRef key="relation"
     minOccurs="1" maxOccurs="1"/>
    <elementRef key="listRelation"
     minOccurs="1" maxOccurs="1"/>
   </alternate>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element listRelation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
```

```
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   ( model.headLike*, desc*, ( model.pLike | ( relation | listRelation )+ ) )
}
```

---

**<listTranspose>** supplies a list of transpositions, each of which is indicated at some point in a document typically by means of metamarks. [*12.3.4.5. Transpositions*]

*Module* transcr — Representation of Primary Sources

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Member of* model.global.meta model.profileDescPart

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal profileDesc scriptNote sponsor

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*core:* desc

*transcr:* transpose

*Example*

```
<listTranspose>
  <transpose>
    <ptr target="#ib02"/>
    <ptr target="#ib01"/>
  </transpose>
</listTranspose>
```

This example might be used for a source document which indicates in some way that the elements identified by `ib02` and code `ib01` should be read in that order (ib02 followed by ib01), rather than in the reading order in which they are presented in the source.

*Content model*

```
<content>
 <sequence>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="transpose" minOccurs="1"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element listTranspose
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( desc*, transpose+ )
}
```

---

**<listWit>** (witness list) lists definitions for all the witnesses referred to by a critical apparatus, optionally grouped hierarchically. [*13.1. The Apparatus Entry, Readings, and Witnesses*]

*Module* textcrit — Critical Apparatus

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.sortable
  - *@sortKey*

*Member of* model.listLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote licence rendition scriptNote sourceDesc tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*textcrit:* lem listWit rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc head

*textcrit:* listWit witness

*Note* May contain a series of <witness> or <listWit> elements.

The provision of a <listWit> element simplifies the automatic processing of the apparatus, e.g. the reconstruction of the readings for all witnesses from an exhaustive apparatus.

Situations commonly arise where there are many more or less fragmentary witnesses, such that there may be quite distinct groups of witnesses for different parts of a text or collection of texts. Such groups may be given separately, or nested within a single <listWit> element at the beginning of the file listing all the witnesses, partial and complete, for the text, with the attestation of fragmentary witnesses indicated within the apparatus by use of the <witStart> and <witEnd> elements described in section *13.1.5. Fragmentary Witnesses*.

Note however that a given witness can only be defined once, and can therefore only appear within a single <listWit> element.

*Example*

```
<listWit>
  <witness  xml:id="HL26">Ellesmere, Huntingdon Library 26.C.9</witness>
  <witness  xml:id="PN392">Hengwrt, National Library of Wales,
     Aberystwyth, Peniarth 392D</witness>
  <witness  xml:id="RP149">Bodleian Library Rawlinson Poetic 149
     (see further <ptr  target="#MSRP149"/>)</witness>
</listWit>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0"/>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="witness"/>
   <elementRef key="listWit"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element listWit
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.sortable.attributes,
   ( model.headLike?, desc*, ( witness | listWit )+ )
}
```

---

**<locale>** contains a brief informal description of the kind of place concerned, for example: a room, a restaurant, a park bench, etc. [*16.2.3. The Setting Description*]

*Module* corpus — Language Corpora

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.settingPart

*Contained by*

*corpus:* setting

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val
*textcrit:* app witDetail
*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data

*Example*

```
<locale>a fashionable restaurant</locale>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element locale
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq.limited}
```

---

## <localProp>

(locally defined property) provides a locally defined character (or glyph) property. [*5.2.1. Character Properties*]

*Module* gaiji — Characters, Glyphs, and Writing Modes

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.gaijiProp
  - *@name*
  - *@value*
  - *@version*
  - *@scheme*

  - att.datable
    * *@period*

    * att.datable.custom
      · *@when-custom*
      · *@notBefore-custom*
      · *@notAfter-custom*
      · *@from-custom*
      · *@to-custom*
      · *@datingPoint*
      · *@datingMethod*
    * att.datable.iso
      · *@when-iso*
      · *@notBefore-iso*
      · *@notAfter-iso*
      · *@from-iso*
      · *@to-iso*
    * att.datable.w3c
      · *@when*
      · *@notBefore*
      · *@notAfter*
      · *@from*
      · *@to*

*Contained by* char glyph

*May contain* Empty element

*Note* No definitive list of local names is proposed. However, the name entity is recommended as a means of naming the property identifying the recommended character entity name for this character or glyph.

*Example*

```
<char xml:id="daikanwaU4EBA">
  <localProp name="name"
    value="CIRCLED IDEOGRAPH 4EBA"/>
  <localProp name="entity" value="daikanwa"/>
  <unicodeProp name="Decomposition_Mapping"
    value="circle"/>
  <mapping type="standard">人</mapping>
</char>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element localProp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.gaijiProp.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   empty
}
```

---

## <location>

(location) defines the location of a place as a set of geographical coordinates, in terms of other named geo-political entities, or as an address. [*14.3.4. Places*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*

  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*

    - *@from*
    - *@to*
  - att.editLike
    - *@evidence*
    - *@instant*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.placeStateLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns place placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*certainty:* precision

*core:* address bibl biblStruct desc email label listBibl measure measureGrp note noteGrp num unit

*header:* biblFull

*msdescription:* depth dim height msDesc width

*namesdates:* affiliation bloc country district geo geogFeat geogName offset placeName region settlement

*Example*

```
<place>
  <placeName>Abbey Dore</placeName>
  <location>
    <geo>51.969604 -2.893146</geo>
  </location>
</place>
```

*Example*

```
<place xml:id="BGbuilding" type="building">
  <placeName>Brasserie Georges</placeName>
  <location>
    <country key="FR"/>
    <settlement type="city">Lyon</settlement>
    <district type="arrondissement">IIème</district>
    <district type="quartier">Perrache</district>
    <placeName type="street">
      <num>30</num>, Cours de Verdun</placeName>
  </location>
</place>
```

*Example*

```
<place type="imaginary">
  <placeName>Atlantis</placeName>
  <location>
    <offset>beyond</offset>
    <placeName>The Pillars of <persName>Hercules</persName>
    </placeName>
  </location>
</place>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <elementRef key="precision"/>
  <classRef key="model.labelLike"/>
  <classRef key="model.placeNamePart"/>
  <classRef key="model.offsetLike"/>
  <classRef key="model.measureLike"/>
  <classRef key="model.addressLike"/>
  <classRef key="model.noteLike"/>
  <classRef key="model.biblLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element location
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
```

```
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.typed.attributes,
   (
      precision    | model.labelLike    | model.placeNamePart    | model.offsetLike    | m
}
```

---

**<locus>** (locus) defines a location within a manuscript, manuscript part, or other object typically as a (possibly discontinuous) sequence of folio references. [*11.3.5. References to Locations within a Manuscript*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*

    - *@resp*
  - att.global.source
    - *@source*
  - att.pointing
    - *@targetLang*
    - *@target*
    - *@evaluate*
  - att.typed
    - *@type*
    - *@subtype*

`@scheme` (scheme) identifies the foliation scheme in terms of which the location is being specified by pointing to some <foliation> element defining it, or to some other equivalent resource.
*Status* Optional
*Datatype* teidata.pointer

`@from` (from) specifies the starting point of the location in a normalized form, typically a page number.
*Status* Optional
*Datatype* teidata.word

`@to` (to) specifies the end-point of the location in a normalized form, typically as a page number.
*Status* Optional
*Datatype* teidata.word

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout locus locusGrp material msItem msItemStruct musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*core:* hi

*gaiji:* g

*msdescription:* locus

character data

*Note* The *target* attribute should only be used to point to elements that contain or indicate a transcription of the locus being described, as in the 'Ben Jonson' example.

To associate a <locus> element with a page image or other comparable representation, the global *facs* attribute should be used, as shown in the 'Birds Praise of Love' example. The *facs* attribute may be used to indicate one or more image files, as in that example, or alternatively it may point to one or more appropriate XML elements, such as the <surface>, <zone>, <graphic>, or <binaryObject> elements.

When a single page is being cited, use the *from* and *to* attributes with an identical value. When no clear endpoint is provided, the *from* attribute may be used without *to*; for example a citation such as 'p. 3ff' might be encoded `<locus from="3">p. 3ff</locus>`.

*Example*

```
<!-- within ms description --><msItem  n="1">
  <locus  target="#F1r #F1v #F2r"  from="1r"
    to="2r">ff. 1r-2r</locus>
  <author>Ben Jonson</author>
  <title>Ode to himself</title>
  <rubric  rend="italics"> An Ode<lb/> to him selfe.</rubric>
  <incipit>Com leaue the loathed stage</incipit>
  <explicit>And see his chariot triumph ore his wayne.</explicit>
  <bibl>
    <name>Beal</name>, <title>Index 1450-1625</title>, JnB 380</bibl>
</msItem>
<!-- within transcription ... -->
<pb  xml:id="F1r"/>
<!-- ... -->
<pb  xml:id="F1v"/>
<!-- ... -->
<pb  xml:id="F2r"/>
<!-- ... -->
```

*Example* The *facs* attribute is available globally when the `transcr` module is included in a schema. It may be used to point directly to an image file, as in the following example:

```
<msItem>
  <locus  facs="images/08v.jpg images/09r.jpg images/09v.jpg images/10r.jpg
images/10v.jpg">fols. 8v-10v</locus>
  <title>Birds Praise of Love</title>
  <bibl>
    <title>IMEV</title>
    <biblScope>1506</biblScope>
  </bibl>
</msItem>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <elementRef key="hi"/>
  <elementRef key="locus"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element locus
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.pointing.attributes,
   att.typed.attributes,
   attribute scheme { text }?,
   attribute from { text }?,
   attribute to { text }?,
   ( text | model.gLike | hi | locus )*
}
```

---

**<locusGrp>** (locus group) groups a number of locations which together form a distinct but discontinuous item within a manuscript, manuscript part, or other object. [*11.3.5. References to Locations within a Manuscript*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@scheme` (scheme) identifies the foliation scheme in terms of which all the locations contained by the group are specified by pointing to some <foliation> element defining it, or to some other equivalent resource.
*Status* Optional
*Datatype* teidata.pointer

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem msItemStruct musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing
*tagdocs:* eg valDesc
*textcrit:* lem rdg wit witDetail witness
*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer
*transcr:* damage fw metamark mod restore retrace secl supplied surplus
*verse:* metSym rhyme
*May contain*
*msdescription:* locus
*Example*

```
<msItem>
  <locusGrp>
    <locus  from="13"  to="26">Bl. 13--26</locus>
    <locus  from="37"  to="58">37--58</locus>
    <locus  from="82"  to="96">82--96</locus>
  </locusGrp>
  <note>Stücke von Daniel Ecklin’s Reise ins h. Land</note>
</msItem>
```

*Content model*

```
<content>
 <elementRef key="locus" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element locusGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute scheme { text }?,
   locus+
}
```

---

**<m>** (morpheme) represents a grammatical morpheme. [*18.1. Linguistic Segment Categories*]
*Module* analysis — Simple Analytic Mechanisms
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.notated
  - *@notation*
- att.segLike
  - *@function*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.fragmentable
    * *@part*
  - att.metrical
    * *@met*
    * *@real*
    * *@rhyme*
- att.typed

- *@type*
- *@subtype*

`@baseForm` supplies the morpheme's base form.
*Status* Optional
*Datatype* teidata.word

*Member of* model.segLike
*Contained by*
*analysis:* cl m phr s w
*cmc:* post
*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castItem role roleDesc sound tech view
*figures:* cell
*header:* change distributor edition extent geoDecl handNote licence scriptNote
*linking:* ab seg
*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark
*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname
*spoken:* u writing
*tagdocs:* eg
*textcrit:* lem rdg wit witDetail
*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer
*transcr:* damage fw metamark mod restore retrace secl supplied surplus
*verse:* rhyme
*May contain*
*analysis:* c interp interpGrp m span spanGrp
*certainty:* certainty precision respons
*core:* cb ellipsis gap gb hi index lb milestone note noteGrp pb q
*figures:* figure notatedMusic
*gaiji:* g
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*spoken:* incident kinesic pause shift vocal writing
*textcrit:* app witDetail
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
character data
*Note* The *type* attribute may be used to indicate the type of morpheme, taking values such as clitic, prefix, stem, etc. as appropriate.
*Example*

```
<w  type="adjective">
  <w  type="noun">
    <m  type="prefix"  baseForm="con">com</m>
    <m  type="root">fort</m>
  </w>
  <m  type="suffix">able</m>
</w>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.hiLike"/>
  <elementRef key="seg"/>
  <elementRef key="m"/>
  <elementRef key="c"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element m
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.notated.attributes,
   att.segLike.attributes,
   att.datcat.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   attribute baseForm { text }?,
   ( text | model.gLike | model.hiLike | seg | m | c | model.global )*
}
```

---

**<macroRef>** points to the specification for some pattern which is to be included in a schema. [*23.7. Macro Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@key` the identifier used for the required pattern within the source indicated.
*Status* Required
*Datatype* teidata.xmlName

*Member of* model.contentPart model.oddRef

*Contained by*

*tagdocs:* alternate content schemaSpec sequence specGrp

*May contain* Empty element

*Note* Patterns or macros are identified by the name supplied as value for the *ident* attribute on the <macroSpec> element in which they are declared. All TEI macro names are unique.

*Example*

```
<schemaSpec ident="myTEI"
  source="http://www.tei-c.org/Vault/P5/current/xml/tei/odd/p5subset.xml">
<!-- ... -->
  <macroRef key="macro.paraContent"/>
<!-- ... -->
</schemaSpec>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element macroRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute key { text },
   empty
}
```

---

## <macroSpec>

<macroSpec> (macro specification) documents the function and implementation of a pattern. [*23.3. Specification Elements 23.7. Macro Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*

- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*
- att.identified
  - *@ident*
  - *@predeclare*
  - *@module*
- att.combinable
  - *@mode*
- att.deprecated
  - *@validUntil*

*Member of* model.oddDecl

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*tagdocs:* schemaSpec specGrp

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* desc gloss

*tagdocs:* altIdent constraintSpec content equiv exemplum listRef remarks valList

*Example*

```
<macroSpec module="tei"
  ident="macro.phraseSeq">
  <content>
    <alternate minOccurs="0"
      maxOccurs="unbounded">
```

```
        <textNode/>
        <classRef  key="model.gLike"/>
        <classRef  key="model.phrase"/>
        <classRef  key="model.global"/>
      </alternate>
    </content>
  </macroSpec>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identSynonyms"/>
   <classRef key="model.descLike"/>
  </alternate>
  <alternate minOccurs="0" maxOccurs="1">
   <elementRef key="content"/>
   <elementRef key="valList"/>
  </alternate>
  <elementRef key="constraintSpec"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="exemplum" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="remarks" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="listRef" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element macroSpec
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.identified.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   (
      ( model.identSynonyms | model.descLike )*,
      ( content | valList )?,
      constraintSpec*,
      exemplum*,
      remarks*,
      listRef*
   )
}
```

**<mapping>** (character mapping) contains one or more characters which are related to the parent character or glyph in some respect, as specified by the *type* attribute. [*5.2. Markup Constructs for Representation of Characters and Glyphs*]

*Module* gaiji — Characters, Glyphs, and Writing Modes

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*

  * *@to-custom*
  * *@datingPoint*
  * *@datingMethod*
- att.datable.iso
  * *@when-iso*
  * *@notBefore-iso*
  * *@notAfter-iso*
  * *@from-iso*
  * *@to-iso*
- att.datable.w3c
  * *@when*
  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*

- att.typed
  - *@type*
  - *@subtype*

*Contained by* *gaiji:* char glyph

*May contain*

*gaiji:* g

character data

*Note* Suggested values for the *type* attribute include exact for exact equivalences, uppercase for uppercase equivalences, lowercase for lowercase equivalences, and simplified for simplified characters. The <g> elements contained by this element can point to either another <char> or <glyph> element or contain a character that is intended to be the target of this mapping.

*Example*

```
<mapping type="modern">r</mapping>
<mapping type="standard">人</mapping>
```

*Content model* `<content> <macroRef key="macro.xtext"/></content>`

*Schema Declaration*

```
element mapping
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
```

```
   att.typed.attributes,
   macro.xtext}
```

---

**<material>** (material) contains a word or phrase describing the material of which the object being described is composed. [*11.3.2. Material and Object Type*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.typed
  - *@type*

– *@subtype*

`@function` describes the function or use of the material in relation to the object as a whole.
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **binding** covering material of a codex
**endband** sewing at the head or tail of the codex spine to strengthen the binding, often decorative
**slipcase** removable protective cover for a set of one or more codices
**support** the surface for writing
**tie** a ribbon or string used to bind or close a codex or rolled scroll
*Note* The sample values here are for descriptive bibliography. Other sets of sample values might include armrests, legs, tabletop, pan, and back for furniture; or wall, floor, window, column, ceiling, roof, stairs, chimney for architecture.

`@target` identifies one or more elements to which the metamark applies.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The *ref* attribute may be used to point to one or more items within a taxonomy of types of material, defined either internally or externally.

*Example*

```
<physDesc>
  <p>
    <material>Parchment</material> leaves with a
  <material>sharkskin</material> binding.</p>
</physDesc>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element material
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   attribute function { text }?,
   attribute target { list { + } }?,
   macro.phraseSeq}
```

---

**<measure>** (measure) contains a word or phrase referring to some quantity of an object or commodity, usually comprising a number, a unit, and a commodity name. [*3.6.3. Numbers and Measures*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*

- att.measurement
  - *@unit*
  - *@unitRef*
  - *@quantity*
  - *@commodity*
- att.ranging
  - *@atLeast*
  - *@atMost*
  - *@min*
  - *@max*
  - *@confidence*
- att.typed
  - ~~type~~
  - @subtype

`@type` specifies the type of measurement in any convenient typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated

*Member of* model.measureLike
*Contained by*
*analysis:* cl phr s span
*cmc:* post
*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure measureGrp meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit
*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castItem role roleDesc sound tech view
*figures:* cell figDesc
*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage
*iso-fs:* fDescr fsDescr
*linking:* ab seg
*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark
*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname
*spoken:* u writing
*tagdocs:* eg valDesc
*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example* This example references a definition of a measurement unit declared in the TEI header:

```
<measure type="weight">
  <num>2</num> pounds of flesh
</measure>
<measure type="currency">£10-11-6d</measure>
<measure type="area" unitRef="#merk">2 <unit>merks</unit> of old
extent</measure>
<!-- In the TEI Header: -->
<encodingDesc>
  <unitDecl>
    <unitDef xml:id="merk" type="area">
      <label>merk</label>
      <placeName ref="#Scotland"/>
      <desc>A merk was an area of land determined variably by its agricultural
            productivity.</desc>
    </unitDef>
  </unitDecl>
</encodingDesc>
```

*Example*

```
<measure quantity="40" unit="hogshead"
  commodity="rum">2 score hh rum</measure>
<measure quantity="12" unit="count"
  commodity="roses">1 doz. roses</measure>
<measure quantity="1" unit="count"
  commodity="tulips">a yellow tulip</measure>
```

*Example*

```
<head>Long papers.</head>
<p>Speakers will be given 30 minutes each: 20 minutes for
presentation, 10 minutes for discussion. Proposals should not
exceed <measure max="500" unit="count"
   commodity="words">500
    words</measure>. This presentation type is suitable for
substantial research, theoretical or critical discussions.</p>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element measure
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.measurement.attributes,
   att.ranging.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.phraseSeq}
```

---

**<measureGrp>** (measure group) contains a group of dimensional specifications which relate to the same object, for example the height and width of a manuscript page. [*11.3.4. Dimensions*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.measurement
  - *@unit*
  - *@unitRef*
  - *@quantity*
  - *@commodity*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.measureLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure measureGrp meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*core:* measure measureGrp num unit

*gaiji:* g

*msdescription:* depth dim height width

*namesdates:* geo

character data

*Example*

```
<measureGrp type="leaves" unit="mm">
  <height scope="range">157-160</height>
  <width quantity="105"/>
</measureGrp>
<measureGrp type="ruledArea" unit="mm">
  <height scope="most" quantity="90"/>
  <width scope="most" quantity="48"/>
</measureGrp>
<measureGrp type="box" unit="in">
  <height quantity="12"/>
  <width quantity="10"/>
  <depth quantity="6"/>
</measureGrp>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.measureLike"/>
```

```
 </alternate>
</content>
```

*Schema Declaration*

```
element measureGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.measurement.attributes,
   att.typed.attributes,
   ( text | model.gLike | model.measureLike )*
}
```

---

**<media>** indicates the location of any form of external media such as an audio or video clip etc. [*3.10. Graphics and Other Non-textual Components*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*
- att.cmc
  - *@generatedBy*
- att.declaring
  - *@decls*
- att.resourced
  - *@url*
- att.timed
  - *@start*
  - *@end*
- att.duration
  - att.duration.iso
    - *@dur-iso*
  - att.duration.w3c
    - *@dur*
- att.typed
  - *@type*
  - *@subtype*
- att.media
  - @width
  - @height
  - @scale

`@mimeType` (MIME media type) specifies the applicable multimedia internet mail extension (MIME) media type.
*Derived from* att.internetMedia
*Status* Required
*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

*Member of* model.graphicLike model.recordingPart
*Contained by*
*analysis:* cl phr s
*cmc:* post
*core:* abbr add addrLine author biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figure formula table

*gaiji:* char glyph

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* recording u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage facsimile fw metamark mod restore retrace secl sourceDoc supplied surface surplus zone

*verse:* rhyme

*May contain*

*core:* desc

*Note* The attributes available for this element are not appropriate in all cases. For example, it makes no sense to specify the temporal duration of a graphic. Such errors are not currently detected.

The *mimeType* attribute must be used to specify the MIME media type of the resource specified by the *url* attribute.

*Example*

```
<figure>
  <media  mimeType="image/png"  url="fig1.png"/>
  <head>Figure One: The View from the Bridge</head>
  <figDesc>A Whistleresque view showing four or five sailing boats in the
foreground, and a
     series of buoys strung out between them.</figDesc>
</figure>
```

*Example*

```
<media  mimeType="audio/wav"
  url="dingDong.wav"  dur="PT10S">
  <desc>Ten seconds of bellringing sound</desc>
</media>
```

*Example*

```
<media  mimeType="video/mp4"
  url="clip45.mp4"  dur="PT45M"  width="500px">
  <desc>A 45 minute video clip to be displayed in a window 500
     px wide</desc>
</media>
```

*Content model*

```
<content>
 <classRef key="model.descLike"
```

```
  minOccurs="0" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element media
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declaring.attributes,
   att.resourced.attributes,
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   att.typed.attributes,
   att.media.attribute.width,
   att.media.attribute.height,
   att.media.attribute.scale,
   attribute mimeType { list { + } },
   model.descLike*
}
```

---

**<meeting>** contains the formalized descriptive title for a meeting or conference, for use in a bibliographic description for an item derived from such a meeting, or as a heading or preamble to publications emanating from it. [*3.12.2.2. Titles, Authors, and Editors*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*

    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*
    - *@to-custom*
    - *@datingPoint*
    - *@datingMethod*
  - att.datable.iso
    - *@when-iso*
    - *@notBefore-iso*
    - *@notAfter-iso*
    - *@from-iso*
    - *@to-iso*
  - att.datable.w3c
    - *@when*
    - *@notBefore*
    - *@notAfter*
    - *@from*
    - *@to*

*Member of* model.divWrapper model.respLike

*Contained by*

*cmc:* post

*core:* bibl lg list monogr

*drama:* castList epilogue performance prologue

*figures:* figure table

*header:* editionStmt titleStmt

*msdescription:* msItem

*textstructure:* body div div1 div2 div3 div4 div5 div6 div7 front group

*May contain*

*core:* abbr address bibl biblStruct choice cit date desc distinct email emph expan foreign gloss hi label list listBibl measure measureGrp mentioned name num ptr q quote ref rs said soCalled stage term time title unit

*dictionaries:* lang

*drama:* camera caption castList move sound tech view

*figures:* table

*header:* biblFull idno

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specGrp specGrpRef tag val

*textcrit:* listApp listWit

*textstructure:* floatingText

*transcr:* am ex subst

character data

*Example*

```
<div>
  <meeting>Ninth International Conference on Middle High German Textual
Criticism, Aachen,
    June 1998.</meeting>
  <list  type="attendance">
    <head>List of Participants</head>
    <item>
      <persName>...</persName>
    </item>
    <item>
      <persName>...</persName>
    </item>
<!--...-->
  </list>
  <p>...</p>
</div>
```

*Content model*

```
<content>
 <macroRef key="macro.limitedContent"/>
```

```
</content>
```

*Schema Declaration*

```
element meeting
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   macro.limitedContent}
```

---

## <memberOf>

**<memberOf>** specifies class membership of the documented element or class. [*23.4.3. Classification of Components*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@key` specifies the identifier for a class of which the documented element or class is a member or subclass.
*Status* Optional
*Datatype* teidata.name

`@mode` specifies the effect of this declaration on its parent module.
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **add** this declaration is added to the current definitions[Default]
**delete** this declaration and all of its children are removed from the current setup

`@max` supplies the maximum number of times the element can occur in elements which use this model class in their content model
*Status* Optional
*Datatype* teidata.numeric

`@min` supplies the minumum number of times the element must occur in elements which use this model class in their content model
*Status* Optional
*Datatype* teidata.numeric

*Contained by* tagdocs: classes
*May contain*
*gaiji:* g
character data

*Note* Elements or classes which are members of multiple (unrelated) classes will have more than one <memberOf> element, grouped by a <classes> element. If an element is a member of a class C1, which is itself a subclass of a class C2, there is no need to state this, other than in the documentation for class C1.

Any additional comment or explanation of the class membership may be provided as content for this element.

*Example*

```
<memberOf key="model.divLike"/>
<memberOf key="att.identified"/>
```

This element will appear in any content model which references `model.divLike`, and will have attributes defined in `att.identified` (in addition to any defined explicitly for this element).

*Content model* `<content> <macroRef key="macro.xtext"/></content>`

*Schema Declaration*

```
element memberOf
{
```

```
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute key { text }?,
   attribute mode { "add" | "delete" }?,
   attribute max { text }?,
   attribute min { text }?,
   macro.xtext}
```

---

**<mentioned>** marks words or phrases mentioned, not used. [*3.3.3. Quotation*]

*Module* core — Elements Available in All TEI Documents

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*

  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*

*Member of* model.emphLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
There is thus a
striking accentual difference between a verbal form like
<mentioned xml:id="X234" xml:lang="el">eluthemen</mentioned>
<gloss target="#X234">we were released,</gloss> accented on the second
syllable of the
word, and its participial derivative
<mentioned xml:id="X235" xml:lang="el">lutheis</mentioned>
<gloss target="#X235">released,</gloss> accented on the last.
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element mentioned
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   macro.phraseSeq}
```

---

## <metamark>

**<metamark>** contains or describes any kind of graphic or written signal within a document the function of which is to determine how it should be read rather than forming part of the actual content of the document. [*12.3.4.2. Metamarks*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.placement
  - *@place*
- att.spanning
  - *@spanTo*

`@function` describes the function (for example status, insertion, deletion, transposition) of the metamark.
*Status* Optional
*Datatype* teidata.word

`@target` identifies one or more elements to which the metamark applies.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Member of* model.global
*Contained by*
*analysis:* cl m phr s span w
*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor ellipsis email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink

objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<surface>
  <metamark  function="used"  rend="line"
    target="#X2"/>
  <zone  xml:id="zone-X2">
    <line>I am that halfgrown <add>angry</add> boy, fallen asleep</line>
    <line>The tears of foolish passion yet undried</line>
    <line>upon my cheeks.</line>
<!-- ... -->
    <line>I pass through <add>the</add> travels and <del>fortunes</del> of
    <retrace>thirty</retrace>
    </line>
    <line>years and become old,</line>
    <line>Each in its due order comes and goes,</line>
    <line>And thus a message for me comes.</line>
    <line>The</line>
  </zone>
  <metamark  function="used"
    target="#zone-X2">Entered - Yes</metamark>
</surface>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element metamark
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.placement.attributes,
   att.spanning.attributes,
   attribute function { text }?,
```

```
   attribute target { list { + } }?,
   macro.specialPara}
```

---

**<metDecl>** (metrical notation declaration) documents the notation employed to represent a metrical pattern when this is specified as the value of a *met*, *real*, or *rhyme* attribute on any structural element of a metrical text (e.g. <lg>, <l>, or <seg>). [*6.6. Metrical Notation Declaration 6.4. Rhyme and Metrical Analysis*]

*Module* verse — Verse

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable
  - *@default*

`@type` indicates whether the notation conveys the abstract metrical form, its actual prosodic realization, or the rhyme scheme, or some combination thereof.

*Status* Optional

*Datatype* 1–3 occurrences of teidata.enumerated separated by whitespace

*Legal values are:* **met** (met attribute) declaration applies to the abstract metrical form recorded on the *met* attribute

**real** (real attribute) declaration applies to the actual realization of the conventional metrical structure recorded on the *real* attribute

**rhyme** declaration applies to the rhyme scheme recorded on the *rhyme* attribute

*Note* By default, the <metDecl> element documents the notation used for metrical pattern and realization. It may also be used to document the notation used for rhyme scheme information; if not otherwise documented, the rhyme scheme notation defaults to the traditional 'abab' notation.

`@pattern` (regular expression pattern) specifies a regular expression defining any value that is legal for this notation.

*Status* Optional

*Datatype* teidata.pattern

*Note* The value must be a valid regular expression per the World Wide Web Consortium's XML Schema Part 2: Datatypes Second Edition, Appendix F.

*Member of* model.encodingDescPart

*Contained by*

*header:* encodingDesc

*May contain*

*core:* note noteGrp p

*linking:* ab

*verse:* metSym

*Note* The encoder may choose whether to define the notation formally or informally. However, the two methods may not be mixed. That is, <metDecl> may contain either a sequence of <metSym> elements or, alternately, a series of paragraphs or other components. If the *pattern* attribute is specified and <metSym> elements are used, then all the codes appearing within the *pattern* attribute should be documented.

Only usable within the header if the verse module is used.

*Example*

```
<metDecl xml:id="ip" type="met"
  pattern="((SU|US)USUSUSUS/)">
  <metSym value="S">stressed syllable</metSym>
  <metSym value="U">unstressed syllable</metSym>
  <metSym value="/">metrical line boundary</metSym>
</metDecl>
```

This example is intended for the far more restricted case typified by the Shakespearean iambic pentameter. Only metrical patterns containing exactly ten syllables, alternately stressed and unstressed, (except for the first two which may be in either order) to each metrical line can be expressed using this notation.

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:metDecl"/> </sch:pattern>

*Content model*

```
<content>
 <alternate>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.pLike"/>
   <classRef key="model.noteLike"/>
  </alternate>
```

```
  <elementRef key="metSym" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element metDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   attribute type
   {
      list
      {
         ( "met" | "real" | "rhyme" ),
         ( "met" | "real" | "rhyme" )?,
         ( "met" | "real" | "rhyme" )?
      }
   }?,
   attribute pattern { text }?,
   ( ( model.pLike | model.noteLike )+ | metSym+ )
}
```

---

## <metSym>

(metrical notation symbol) documents the intended significance of a particular character or character sequence within a metrical notation, either explicitly or in terms of other <metSym> elements in the same <metDecl>. [*6.6. Metrical Notation Declaration*]

*Module* verse — Verse

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@value` specifies the character or character sequence being documented.
*Status* Required
*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

`@terminal` specifies whether the symbol is defined in terms of other symbols (*terminal* is set to false) or in prose (*terminal* is set to true).
*Status* Optional
*Datatype* teidata.truthValue
*Default* true
*Note* The value true indicates that the element contains a prose definition of its meaning; the value false indicates that the element contains a definition of its meaning given using symbols defined elsewhere in the same <metDecl> element.

*Contained by:* metDecl
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit
*dictionaries:* lang
*figures:* figure notatedMusic
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val
*textcrit:* app witDetail
*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data
*Example*

```
<metSym value="x">a stressed syllable</metSym>
<metSym value="o">an unstressed syllable</metSym>
<metSym value="A" terminal="false">xoo</metSym>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element metSym
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute value { list { + } },
   attribute terminal { text }?,
   macro.phraseSeq.limited}
```

---

## <milestone>

(milestone) marks a boundary point separating any kind of section of a text, typically but not necessarily indicating a point at which some part of a standard reference system changes, where the change is not represented by a structural element. [*3.11.3. Milestone Elements*]

*Module* core — Elements Available in All TEI Documents
*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*

- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.breaking
  - *@break*
- att.cmc
  - *@generatedBy*
- att.edition
  - *@ed*
  - *@edRef*
- att.milestoneUnit
  - *@unit*
- att.spanning
  - *@spanTo*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.milestoneLike

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list listBibl measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc subst supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Note* For this element, the global *n* attribute indicates the new number or other value for the unit which changes at this milestone. The special value *unnumbered* should be used in passages which fall outside the normal numbering scheme, such as chapter or other headings, poem numbers or titles, etc.

The order in which <milestone> elements are given at a given point is not normally significant.

*Example*

```
<milestone n="23" ed="La" unit="Dreissiger"/>
... <milestone n="24" ed="AV" unit="verse"/> ...
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element milestone
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.breaking.attributes,
   att.cmc.attributes,
   att.edition.attributes,
   att.milestoneUnit.attributes,
   att.spanning.attributes,
   att.typed.attributes,
   empty
}
```

**<mod>** represents any kind of modification identified within a single document. [*12.3.4.1. Generic Modification*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*
  - att.ranging
    - *@atLeast*

    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
  - att.spanning
    - *@spanTo*
  - att.transcriptional
    - *@status*
    - *@cause*
    - *@seq*
    - att.editLike
    * *@evidence*
    * *@instant*
    - att.placement
    * *@place*
    - att.written
    * *@hand*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.linePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<mod  type="subst">
  <add>pleasing</add>
  <del>agreable</del>
</mod>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element mod
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.spanning.attributes,
   att.transcriptional.attributes,
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   att.typed.attributes,
   macro.paraContent}
```

---

**<model>** describes the processing intended for a specified element. [*23.5.4.1. The TEI processing model*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

- att.predicate
  - *@predicate*

`@behaviour` names the process or function which this processing model uses in order to produce output.
*Status* Required
*Datatype* teidata.enumerated
*Suggested values include:* **alternate** (default, alternate) support display of alternative visualisations, for example by displaying the preferred content, by displaying both in parallel, or by toggling between the two.
**anchor** (id) create an identifiable anchor point in the output.
**block** (content) create a block structure
**body** (content) create the body of a document.
**break** (type, label) create a line, column, or page break according to the value of `type`
**cell** (content) create a table cell
**cit** (content, source) show the content, with an indication of the source
**document** (content) start a new output document
**figure** (title) make a figure with the *title* as caption
**glyph** (uri) show a character by looking up reference to a chardesc at the given URI
**graphic** (url, width, height, scale, title) if `url` is present, uses it to display graphic, else display a placeholder image.
**heading** (content, level) creates a heading.
**index** (type) generate list according to `type`.
**inline** (content, label) creates inline element out of `content`
**link** (content, uri) create hyperlink
**list** (content) create a list.
**listItem** (content) create a list item.
**metadata** (content) create metadata section
**note** (content, place, label) create a note, often out of line, depending on the value of `place`; could be margin, footnote, endnote, inline
**omit** do nothing, do not process children
**paragraph** (content) create a paragraph out of `content`.
**row** (content) create a table row
**section** (content) create a new section of the output document
**table** (content) create a table
**text** (content) create literal text
**title** (content) create document title

`@useSourceRendition` whether to obey any rendition attribute that is present.
*Status* Optional
*Datatype* teidata.truthValue
*Note* If the *useSourceRendition* attribute is not specified, or has the value false, any renditional information specified in the source document should be ignored.

`@output` the intended output.

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **web** the output is intended for presentation in a web format

**print** the output is intended for presentation in a print format

**plain** the output is intended for presentation in a plain text format

*Note* If the *output* attribute is not specified, this model is assumed to apply for all forms of output.

`@cssClass` the name of a CSS class which should be associated with this element

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.name separated by whitespace

*Contained by tagdocs:* elementSpec modelGrp modelSequence

*May contain*

*core:* desc gloss

*tagdocs:* equiv outputRendition param

*Note* The intended output to be generated for a particular behaviour of a processing model may be documented in one or all of the three following ways. Firstly, the *cssClass* attribute may be used to specify the name of a CSS style in some associated CSS stylesheet which is to be applied to each occurrence of a specified element found (in a given context, for a specified output). Secondly, the attribute *useSourceRendition* may be used to indicate that the default rendition for occurrences of this element, as defined by a <rendition> element in the document header, should be applied. Thirdly, the styling to be applied may be specified explicitly as content of a child <outputRendition> element. When more than one of these options is used, they are understood to be combined in accordance with the rules for multiple declaration of the styling language used.

*Example*

```
<model behaviour="inline"/>
```

*Example*

```
<elementSpec mode="change" ident="quote">
  <model predicate="ancestor::p"
    behaviour="inline">
    <desc versionDate="2015-08-21"
      xml:lang="en">If it's inside a paragraph then it's
        inline</desc>
  </model>
  <model predicate="not(ancestor::p)"
    behaviour="block">
    <desc versionDate="2015-08-21"
      xml:lang="en">If it's not inside a paragraph then it is
        block level</desc>
  </model>
</elementSpec>
```

*Example*

```
<model predicate="parent::person"
  behaviour="inline">
  <desc versionDate="2015-08-21"
    xml:lang="en">If it is a child of a person element, treat as
     inline</desc>
</model>
```

*Example*

```
<model behaviour="inline">
  <outputRendition>font-weight:bold</outputRendition>
</model>
```

*Example*

```
<model behaviour="inline" output="print"/>
```

*Example*

```
<model predicate="ancestor::p"
  behaviour="inline">
  <param name="content" value="@n"/>
</model>
```

*Example*

```
<elementSpec mode="change" ident="date">
  <model output="print" predicate="text()"
    behaviour="inline"/>
  <model output="print"
    predicate="@when and not(text())" behaviour="inline">
    <param name="content" value="@when"/>
  </model>
  <model predicate="@when" output="web"
    behaviour="alternate">
    <param name="alternate" value="@when"/>
  </model>
</elementSpec>
```

*Example*

```
<elementSpec mode="change" ident="choice">
  <model predicate="sic and corr"
    behaviour="alternate">
    <param name="default" value="corr"/>
    <param name="alternate" value="sic"/>
  </model>
  <model predicate="abbr and expan"
    behaviour="alternate">
    <param name="default" value="expan"/>
    <param name="alternate" value="abbr"/>
  </model>
  <model predicate="orig and reg"
    behaviour="alternate">
    <param name="default" value="reg"/>
    <param name="alternate" value="orig"/>
  </model>
</elementSpec>
```

*Schematron* It is ambigous to have 2 <model>s that have both the same *output* and have no *predicate* (and thus select the same set of nodes). <sch:rule context="tei:model[ not( parent::tei:modelSequence ) ][ not( @predicate ) ]"> <sch:let name="output" value="normalize-space( @output )"/> <sch:report test="following-sibling::tei:model [ not( @predicate )] [ normalize-space( @output ) eq $output ]"> There are 2 (or more) 'model' elements in this '<sch:value-of select="local-name(..)"/>' that

have no predicate, but are targeted to the same output (”<sch:value-of select=”( $output, parent::modelGrp/@output, ’all’)[1]”/>”)</sch:report> </sch:rule>

*Schematron* It is ambigous to have 2 <model>s that have both the same *output* and the same *predicate* (and thus select the same set of nodes). <sch:rule context=”tei:model[ not( parent::tei:modelSequence ) ][ @predicate ]”> <sch:let name=”predicate” value=”normalize-space( @predicate )”/> <sch:let name=”output” value=”normalize-space( @output )”/> <sch:report test=”following-sibling::tei:model [ normalize-space( @predicate ) eq $predicate ] [ normalize-space( @output ) eq $output ]”> There are 2 (or more) ’model’ elements in this ’<sch:value-of select=”local-name(..)”/>’ that have the same predicate, and are targeted to the same output (”<sch:value-of select=”( $output, parent::modelGrp/@output, ’all’)[1]”/>”)</sch:report> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identEquiv"/>
   <classRef key="model.descLike"/>
  </alternate>
  <elementRef key="param" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="outputRendition"
   minOccurs="0" maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element model
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.predicate.attributes,
   attribute behaviour
   {
      "alternate"
    | "anchor"
    | "block"
    | "body"
    | "break"
    | "cell"
    | "cit"
    | "document"
    | "figure"
    | "glyph"
    | "graphic"
    | "heading"
    | "index"
    | "inline"
    | "link"
    | "list"
```

```
    | "listItem"
    | "metadata"
    | "note"
    | "omit"
    | "paragraph"
    | "row"
    | "section"
    | "table"
    | "text"
    | "title"
   },
   attribute useSourceRendition { text }?,
   attribute output { text }?,
   attribute cssClass { list { + } }?,
   ( ( model.identEquiv | model.descLike )*, param*, outputRendition* )
}
```

---

**<modelGrp>** (model group) any grouping of <model> or <modelSequence> elements with a common output method. [*23.5.4.4. Model Contexts and Outputs*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*
– att.global.source
* *@source*

`@useSourceRendition` whether to obey any rendition attribute which is present.
*Status* Optional
*Datatype* teidata.truthValue

`@output` the intended output method.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **web** the output is intended for presentation in a web format
**print** the output is intended for presentation in a print format
**plaintext** the output is intended for presentation in a plain text format

*Contained by tagdocs:* elementSpec
*May contain*
*core:* desc gloss
*tagdocs:* equiv model modelSequence outputRendition
*Note* The child <model> elements of a <modelGrp> are always processed independently.
*Example*

```
<elementSpec mode="change" ident="abbr">
  <modelGrp output="web">
    <model predicate="parent::choice"
      behaviour="omit"/>
    <model predicate="ancestor::front"
      behaviour="inline">
      <outputRendition>font-style:italic; </outputRendition>
    </model>
    <model predicate="not(parent::choice)"
      behaviour="inline">
      <outputRendition scope="before">content: ' ('</outputRendition>
      <outputRendition scope="after">content: ')'</outputRendition>
    </model>
  </modelGrp>
  <modelGrp output="print">
    <model predicate="parent::choice"
      behaviour="omit"/>
    <model predicate="not(parent::choice)"
      behaviour="note">
      <param name="place" value="'foot'"/>
    </model>
  </modelGrp>
</elementSpec>
```

*Example*

```
<modelGrp output="print">
  <modelSequence>
    <model behaviour="inline">
      <param name="content" value="@n"/>
```

```
      </model>
      <model  behaviour="note">
        <param  name="place"  value="'foot'"/>
      </model>
    </modelSequence>
  </modelGrp>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identEquiv"/>
   <classRef key="model.descLike"/>
  </alternate>
  <elementRef key="outputRendition"
   minOccurs="0"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="modelSequence"/>
   <elementRef key="model"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element modelGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute useSourceRendition { text }?,
   attribute output { "web" | "print" | "plaintext" }?,
   (
      ( model.identEquiv | model.descLike )*,
      outputRendition?,
      ( modelSequence | model )+
   )
}
```

---

**<modelSequence>** any sequence of model or <modelSequence> elements which is to be processed as a single set of actions. [*23.5.4.7. Model sequence*]

*Module* tagdocs — Documentation Elements

*Attributes* • att.global

- *@xml:id*
- *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@predicate` the condition under which this model applies given as an XPath Predicate Expression.
*Status* Optional
*Datatype* teidata.xpath

`@useSourceRendition` whether to obey any rendition attribute that is present.
*Status* Optional
*Datatype* teidata.truthValue

`@output` the intended output method.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **web** the output is intended for presentation in a web format
**print** the output is intended for presentation in a print format
**plaintext** the output is intended for presentation in a plain text format

*Contained by:* tagdocs: elementSpec modelGrp
*May contain*

*core:* desc gloss

*tagdocs:* equiv model

*Note* All of the <model> elements inside a <modelSequence> element must be processed. The operations indicated by each <model> child cannot be dependent on any other operation in the sequence. The results must be rendered in the order given by the list of <model>s in the <modelSequence>.

*Example*

```
<modelGrp output="print">
  <modelSequence>
    <model behaviour="inline">
      <param name="content" value="@n"/>
    </model>
    <model behaviour="footnote">
      <param name="place" value="'foot'"/>
    </model>
  </modelSequence>
</modelGrp>
```

*Schematron* <sch:rule context="tei:modelSequence"> <sch:report test="tei:model[@output]" role="warning">The 'model' children of a 'modelSequence' element inherit the @output attribute of the parent 'modelSequence', and thus should not have their own</sch:report> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identEquiv"/>
   <classRef key="model.descLike"/>
  </alternate>
  <elementRef key="model" minOccurs="2"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element modelSequence
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute predicate { text }?,
   attribute useSourceRendition { text }?,
   attribute output { "web" | "print" | "plaintext" }?,
   ( ( model.identEquiv | model.descLike )*, ( model, model, model* ) )
}
```

**<moduleRef>** (module reference) references a module which is to be incorporated into a schema. [*23.2. Modules and Schemas*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

`@prefix` specifies a default prefix which will be prepended to all patterns from the imported module.

*Status* Optional

*Datatype* 0–1 occurrences of teidata.xmlName separated by whitespace

*Schematron* <sch:rule context="tei:moduleRef"> <sch:report test="//*[ not( generate-id(.) eq generate-id( current() ) ) ]/@prefix = @prefix">The prefix attribute of <sch:name/> should not match that of any other element (it would defeat the purpose)</sch:report> </sch:rule>

*Schematron* <sch:rule context="tei:moduleRef"> <sch:report test="@except and @include">It is an error to supply both the @include and @except attributes</sch:report> </sch:rule>

*Note* Use of this attribute avoids name collisions (and thus invalid schemas) when the external schema being mixed in with TEI uses a name the TEI or some other included external schema already uses for a pattern.

`@include` supplies a list of the elements which are to be copied from the specified module into the schema being defined.
*Status* Optional
*Datatype* 0–∞ occurrences of teidata.xmlName separated by whitespace

`@except` supplies a list of the elements which are not to be copied from the specified module into the schema being defined.
*Status* Optional
*Datatype* 0–∞ occurrences of teidata.xmlName separated by whitespace

`@key` the name of a TEI module
*Status* Optional
*Datatype* teidata.xmlName

`@url` (uniform resource locator) refers to a non-TEI module of RELAX NG code by external location
*Status* Optional
*Datatype* teidata.pointer

*Member of* model.oddRef
*Contained by*
*tagdocs:* schemaSpec specGrp
*May contain*
*tagdocs:* content

*Note* If neither *include* nor *except* is supplied, the effect of this element is to make all the declarations contained by the referenced module available to the schema being compiled. If both attributes are supplied, an ODD processor should signal an error.

A TEI module is identified by the name supplied as value for the *ident* attribute on a <moduleSpec> element. The *source* attribute may be used to specify an online source from which the specification of that module may be read. A URI may alternatively be supplied in the case of a non-TEI module, and this is expected to be written as a RELAX NG schema.

If the *url* attribute is used, the <content> element may also be supplied as a child of this element. Its content (which is assumed to be a fragment of RELAX NG code) will be copied along with the content of the resource indicated by the *url* attribute into the target RELAX NG schema.

*Example*

```
<moduleRef key="linking"/>
```

This includes all objects available from the linking module.

*Example*

```
<moduleRef key="linking"
  except="linkGrp link"/>
```

This includes all elements available from the linking module except for the <link> and <linkGrp> elements.

*Example*

```
<moduleRef key="linking"
  include="linkGrp link"/>
```

This includes only the <link> and <linkGrp> elements from the linking module.

*Schematron* <sch:rule context="tei:moduleRef"> <sch:report test="* and @key"> Child elements of <sch:name/> are only allowed when an external module is being loaded </sch:report> </sch:rule>

*Content model*

```
<content>
 <elementRef key="content" minOccurs="0"/>
</content>
```

*Schema Declaration*

```
element moduleRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute prefix { ? }?,
   ( attribute include { list { * } }? | attribute except { list { * } }? ),
   ( attribute key { text }? | attribute url { text }? ),
   content?
}
```

---

**<moduleSpec>** (module specification) documents the structure, content, and purpose of a single module, i.e. a named and externally visible group of declarations. [*23.2. Modules and Schemas*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.identified
  - *@ident*
  - *@predeclare*
  - *@module*
  - att.combinable
    * *@mode*
    * att.deprecated
    · *@validUntil*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.oddDecl

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*tagdocs:* schemaSpec specGrp

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus
*verse:* rhyme
*May contain*
*core:* desc gloss
*header:* idno
*tagdocs:* equiv exemplum listRef remarks
*Example*

```
<moduleSpec  ident="namesdates">
  <idno  type="FPI">Names and Dates</idno>
  <desc>Additional elements for names and dates</desc>
</moduleSpec>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identEquiv"/>
   <elementRef key="idno"/>
   <classRef key="model.descLike"/>
  </alternate>
  <elementRef key="exemplum" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="remarks" minOccurs="0"/>
  <elementRef key="listRef" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element moduleSpec
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.identified.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   att.typed.attributes,
   (
      ( model.identEquiv | idno | model.descLike )*,
      exemplum*,
      remarks?,
      listRef*
   )
}
```

**<monogr>** (monographic level) contains bibliographic elements describing an item (e.g. a book or journal) published as an independent item (i.e. as a separate physical object). [*3.12.2.1. Analytic, Monographic, and Series Levels*]

*Module* core — Elements Available in All TEI Documents

*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Contained by* biblStruct

*May contain*

*core:* author biblScope editor imprint meeting note noteGrp ptr ref respStmt textLang title

*header:* authority availability edition extent funder idno sponsor

*tagdocs:* listRef

*Note* May contain specialized bibliographic elements, in a prescribed order.

The <monogr> element may only occur only within a <biblStruct>, where its use is mandatory for the description of a monographic-level bibliographic item.

*Example*

```
<biblStruct>
  <analytic>
    <author>Chesnutt, David</author>
    <title>Historical Editions in the States</title>
  </analytic>
  <monogr>
    <title  level="j">Computers and the Humanities</title>
    <imprint>
      <date  when="1991-12">(December, 1991):</date>
    </imprint>
    <biblScope>25.6</biblScope>
    <biblScope  unit="page"  from="377"  to="380">377–380</biblScope>
  </monogr>
</biblStruct>
```

*Example*

```
<biblStruct  type="book">
  <monogr>
    <author>
      <persName>
        <forename>Leo Joachim</forename>
        <surname>Frachtenberg</surname>
      </persName>
    </author>
    <title  type="main"  level="m">Lower Umpqua Texts</title>
    <imprint>
      <pubPlace>New York</pubPlace>
      <publisher>Columbia University Press</publisher>
      <date>1914</date>
    </imprint>
  </monogr>
  <series>
    <title  type="main"  level="s">Columbia University Contributions to
        Anthropology</title>
    <biblScope  unit="volume">4</biblScope>
  </series>
</biblStruct>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0">
   <sequence>
    <alternate>
     <elementRef key="author"/>
     <elementRef key="editor"/>
     <elementRef key="meeting"/>
     <elementRef key="respStmt"/>
    </alternate>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <elementRef key="author"/>
     <elementRef key="editor"/>
     <elementRef key="meeting"/>
     <elementRef key="respStmt"/>
    </alternate>
```

```
    <elementRef key="title" minOccurs="1"
     maxOccurs="unbounded"/>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <classRef key="model.ptrLike"/>
     <elementRef key="idno"/>
     <elementRef key="textLang"/>
     <elementRef key="editor"/>
     <elementRef key="respStmt"/>
    </alternate>
   </sequence>
   <sequence>
    <alternate minOccurs="1"
     maxOccurs="unbounded">
     <elementRef key="title"/>
     <classRef key="model.ptrLike"/>
     <elementRef key="idno"/>
    </alternate>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <elementRef key="textLang"/>
     <elementRef key="author"/>
     <elementRef key="editor"/>
     <elementRef key="meeting"/>
     <elementRef key="respStmt"/>
    </alternate>
   </sequence>
   <sequence>
    <elementRef key="authority"/>
    <elementRef key="idno"/>
   </sequence>
  </alternate>
  <elementRef key="availability"
   minOccurs="0" maxOccurs="unbounded"/>
  <classRef key="model.noteLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="edition"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="idno"/>
    <classRef key="model.ptrLike"/>
    <elementRef key="editor"/>
    <elementRef key="sponsor"/>
    <elementRef key="funder"/>
    <elementRef key="respStmt"/>
   </alternate>
  </sequence>
  <elementRef key="imprint"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="imprint"/>
   <elementRef key="extent"/>
   <elementRef key="biblScope"/>
  </alternate>
 </sequence>
```

```
</content>
```

*Schema Declaration*

```
element monogr
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      (
         (
            ( author | editor | meeting | respStmt ),
            ( author | editor | meeting | respStmt )*,
            title+,
            ( model.ptrLike | idno | textLang | editor | respStmt )*
         )
       | (
            ( title | model.ptrLike | idno )+,
            ( textLang | author | editor | meeting | respStmt )*
         )
       | ( authority, idno )
      )?,
      availability*,
      model.noteLike*,
      (
         (
            edition,
            ( idno | model.ptrLike | editor | sponsor | funder | respStmt )*
         )*
      ),
      imprint,
      ( imprint | extent | biblScope )*
   )
}
```

---

**<mood>** (mood) contains information about the grammatical mood of verbs (e.g. indicative, subjunctive, imperative). [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.lexicographic.normalized
    * *@norm*
    * *@orig*

*Member of* model.morphLike
*Contained by*
*dictionaries:* dictScrap entryFree form gramGrp
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* This element is synonymous with <gram type="mood">.

*Example* Taken from *Wörterbuch der Deutschen Sprache. Veranstaltet und herausgegeben von Joachim Heinrich Campe. Vierter Theil. S - bis - T. (Braunschweig 1810. In der Schulbuchhandlung)*:

> Treffen, v. unregelm. … du triffst, …

```
<entry>
  <form type="inflected">
    <gramGrp>
      <per value="2"/>
      <number value="singular"/>
      <tns value="present"/>
      <mood value="indicative"/>
    </gramGrp>
    <form type="personalpronoun">
      <orth>du</orth>
    </form>
    <form type="headword">
      <orth>
        <oRef>triffst</oRef>
      </orth>
    </form>
  </form>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element mood
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   macro.paraContent}
```

---

**<move>** (movement) marks the actual movement of one or more characters. [*7.2.4. Stage Directions*]

*Module* drama — Performance Texts

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.ascribed.directed
  - *@toWhom*

- att.ascribed
  - *@who*

- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the movement, for example as an entrance or exit.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **entrance** character is entering the stage.

**exit** character is exiting the stage.

**onStage** character moves on stage

`@where` specifies the direction of a stage movement.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.authority separated by whitespace

*Sample values include:* **L** (left) stage left

**R** (right) stage right

**C** (center) centre stage

*Note* Full blocking information will normally require combinations of values, (for example 'UL' for 'upper stage left') and may also require more detailed encoding of speed, direction etc. Full documentation of any coding system used should be provided in the header. URIs may be used as values.

`@perf` (performance) identifies the performance or performances in which this movement occurred as specified by pointing to one or more <performance> elements.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Member of* model.stageLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l lg meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound spGrp tech view

*figures:* cell figDesc figure
*header:* change handNote licence rendition scriptNote tagUsage
*iso-fs:* fDescr fsDescr
*linking:* ab seg
*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote
*namesdates:* occupation
*spoken:* u writing
*textcrit:* lem rdg witness
*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer
*transcr:* damage metamark mod restore retrace secl supplied surplus
*verse:* rhyme
*May contain* Empty element
*Example*

```
<performance xml:id="perf1">
  <p>First performance</p>
  <castList>
    <castItem>
      <role xml:id="bellaf">Bellafront</role>
    </castItem>
<!-- ... -->
  </castList>
</performance>
<!-- ... -->
<stage type="entrance">
  <move who="#bellaf" type="enter" where="L"
    perf="#perf1"/> Enter
Bellafront mad.
</stage>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element move
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.directed.attributes,
   att.ascribed.attributes,
   att.typed.attribute.subtype,
   attribute type { "entrance" | "exit" | "onStage" }?,
   attribute where { list { + } }?,
   attribute perf { list { + } }?,
   empty
}
```

---

**<msContents>** (manuscript contents) describes the intellectual content of a manuscript, manuscript part, or other object either as a series of paragraphs or as a series of structured manuscript items. [*11.6. Intellectual Content*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.msClass
  - *@class*
- att.msExcerpt
  - *@defective*

*Contained by:* msDescription: msDesc msFrag msPart
*namesdates:* object

*May contain*

*core:* p textLang

*linking:* ab

*msdescription:* msItem msItemStruct summary

*textstructure:* titlePage

*Note* Unless it contains a simple prose description, this element should contain at least one of the elements <summary>, <msItem>, or <msItemStruct>. This constraint is not currently enforced by the schema.

*Example*

```
<msContents class="#sermons">
  <p>A collection of Lollard sermons</p>
</msContents>
```

*Example*

```
<msContents>
  <msItem n="1">
    <locus>fols. 5r-7v</locus>
    <title>An ABC</title>
    <bibl>
      <title>IMEV</title>
      <biblScope>239</biblScope>
    </bibl>
  </msItem>
  <msItem n="2">
    <locus>fols. 7v-8v</locus>
    <title xml:lang="frm">Lenvoy de Chaucer a Scogan</title>
    <bibl>
      <title>IMEV</title>
      <biblScope>3747</biblScope>
    </bibl>
  </msItem>
  <msItem n="3">
    <locus>fol. 8v</locus>
    <title>Truth</title>
    <bibl>
      <title>IMEV</title>
      <biblScope>809</biblScope>
    </bibl>
  </msItem>
  <msItem n="4">
    <locus>fols. 8v-10v</locus>
    <title>Birds Praise of Love</title>
    <bibl>
      <title>IMEV</title>
      <biblScope>1506</biblScope>
    </bibl>
  </msItem>
  <msItem n="5">
    <locus>fols. 10v-11v</locus>
    <title xml:lang="la">De amico ad amicam</title>
    <title xml:lang="la">Responcio</title>
    <bibl>
      <title>IMEV</title>
      <biblScope>16 & 19</biblScope>
    </bibl>
```

```
  </msItem>
  <msItem  n="6">
    <locus>fols. 14r-126v</locus>
    <title>Troilus and Criseyde</title>
    <note>Bk. 1:71-Bk. 5:1701, with additional losses due to mutilation
throughout</note>
  </msItem>
</msContents>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="summary" minOccurs="0"/>
   <elementRef key="textLang" minOccurs="0"/>
   <elementRef key="titlePage"
    minOccurs="0"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="msItem"/>
    <elementRef key="msItemStruct"/>
   </alternate>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element msContents
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.msClass.attributes,
   att.msExcerpt.attributes,
   (
      model.pLike+
    | ( summary?, textLang?, titlePage?, ( msItem | msItemStruct )* )
   )
}
```

---

## <msDesc>

(manuscript description) contains a description of a single identifiable manuscript or other text-bearing object such as an early printed book. [*11.1. Overview*]

*Module* msdescription — Manuscript Description
*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.declaring
  - *@decls*
- att.docStatus
  - *@status*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.biblLike

*Contained by*

*cmc:* post

*core:* add cit corr del desc emph head hi item l listBibl meeting note orig p q quote ref reg relatedItem said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote sourceDesc tagUsage taxonomy

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout msItem musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* climate event location object occupation org person personGrp persona place population state terrain trait

*spoken:* broadcast scriptStmt u writing

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* head p

*linking:* ab

*msdescription:* additional history msContents msFrag msIdentifier msPart physDesc

*Note* Although the <msDesc> has primarily been designed with a view to encoding manuscript descriptions, it may also be used for other objects such as early printed books, fascicles, epigraphs, or any text-bearing objects that require substantial description. If an object is not text-bearing or the reasons for describing the object is not primarily the textual content, the more general <object> may be more suitable.

*Example*

```
<msDesc>
  <msIdentifier>
    <settlement>Oxford</settlement>
    <repository>Bodleian Library</repository>
    <idno  type="Bod">MS Poet. Rawl. D. 169.</idno>
  </msIdentifier>
  <msContents>
    <msItem>
      <author>Geoffrey Chaucer</author>
      <title>The Canterbury Tales</title>
    </msItem>
  </msContents>
  <physDesc>
    <objectDesc>
      <p>A parchment codex of 136 folios, measuring approx
            28 by 19 inches, and containing 24 quires.</p>
      <p>The pages are margined and ruled throughout.</p>
      <p>Four hands have been identified in the manuscript: the first 44
            folios being written in two cursive anglicana scripts, while the
            remainder is for the most part in a mixed secretary hand.</p>
    </objectDesc>
  </physDesc>
</msDesc>
```

*Schematron* <sch:rule context="tei:msContents|tei:physDesc|tei:history|tei:additional"> <!– Note: This rule applies to <msContents>, <physDesc>, <history>, and <additional> wherever they occur. Luckily they

are only allowed in places where they are constrained to 0 or 1 occurrence. If that changes someday, this constraint may will likely need to be updated, too. –> <!– Also worth noting that a) if & when we can use abstract patterns, this would be better handled as a single abstract rule somewhere, and concrete rules in the individual <elementSpec>s; and b) I did not test for the existence of "../*[name(.) eq $gi][2]" because then an error would be generated for each of the multiple occurrences of $gi. –> <sch:let name="gi" value="name(.)"/> <sch:report test="preceding-sibling::*[ name(.) eq $gi ] and not( following-sibling::*[ name(.) eq $gi ] )"> Only one <sch:name/> is allowed as a child of <sch:value-of select="name(..)"/>. </sch:report> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <elementRef key="msIdentifier"/>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate>
   <classRef key="model.pLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="msContents"/>
    <elementRef key="physDesc"/>
    <elementRef key="history"/>
    <elementRef key="additional"/>
    <elementRef key="msPart"/>
    <elementRef key="msFrag"/>
   </alternate>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element msDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.docStatus.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      msIdentifier,
      model.headLike*,
      (
         model.pLike+
       | ( msContents | physDesc | history | additional | msPart | msFrag )*
      )
   )
}
```

---

**<msFrag>** (manuscript fragment) contains information about a fragment described in relation to a prior context, typically as a description of a virtual reconstruction of a manuscript or other object whose fragments were catalogued separately. [*11.11. Manuscript Fragments*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - *@type*
  - *@subtype*

*Contained by* *msDescription:* msDesc

*May contain*

*core:* head p

*linking:* ab

*msdescription:* additional altIdentifier history msContents msIdentifier physDesc
*Example*

```
<msDesc>
  <msIdentifier>
    <msName  xml:lang="la">Codex Suprasliensis</msName>
  </msIdentifier>
  <msFrag>
    <msIdentifier>
      <settlement>Ljubljana</settlement>
      <repository>Narodna in univerzitetna knjiznica</repository>
      <idno>MS Kopitar 2</idno>
    </msIdentifier>
    <msContents>
      <summary>Contains ff. 10 to 42 only</summary>
    </msContents>
  </msFrag>
  <msFrag>
    <msIdentifier>
      <settlement>Warszawa</settlement>
      <repository>Biblioteka Narodowa</repository>
      <idno>BO 3.201</idno>
    </msIdentifier>
  </msFrag>
  <msFrag>
    <msIdentifier>
      <settlement>Sankt-Peterburg</settlement>
      <repository>Rossiiskaia natsional'naia biblioteka</repository>
      <idno>Q.p.I.72</idno>
    </msIdentifier>
  </msFrag>
</msDesc>
```

*Content model*

```
<content>
 <sequence>
  <alternate>
   <elementRef key="altIdentifier"/>
   <elementRef key="msIdentifier"/>
  </alternate>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate>
   <classRef key="model.pLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="msContents"/>
    <elementRef key="physDesc"/>
    <elementRef key="history"/>
    <elementRef key="additional"/>
   </alternate>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element msFrag
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   (
      ( altIdentifier | msIdentifier ),
      model.headLike*,
      ( model.pLike+ | ( msContents | physDesc | history | additional )* )
   )
}
```

---

## <msIdentifier>

(manuscript identifier) contains the information required to identify the manuscript or similar object being described. [*11.4. The Manuscript Identifier*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*

* *@rendition*

– att.global.responsibility

* *@cert*
* *@resp*

– att.global.source

* *@source*

*Member of* model.biblPart

*Contained by*

*core:* bibl

*msdescription:* msDesc msFrag msPart

*May contain*

*header:* idno

*msdescription:* altIdentifier collection institution msName repository

*namesdates:* bloc country district geogName objectName placeName region settlement

*Example*

```
<msIdentifier>
  <settlement>San Marino</settlement>
  <repository>Huntington Library</repository>
  <idno>MS.El.26.C.9</idno>
</msIdentifier>
```

*Schematron* <sch:rule context="tei:msIdentifier"> <sch:report test="not( parent::tei:msPart ) and ( child::*[1]/self::idno or child::*[1]/self::altIdentifier or normalize-space(.) eq '')">An msIdentifier must contain either a repository or location.</sch:report> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <sequence>
   <classRef key="model.placeNamePart"
    expand="sequenceOptional"/>
   <elementRef key="institution"
    minOccurs="0"/>
   <elementRef key="repository"
    minOccurs="0"/>
   <elementRef key="collection"
    minOccurs="0" maxOccurs="unbounded"/>
   <elementRef key="idno" minOccurs="0"
    maxOccurs="unbounded"/>
  </sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="msName"/>
   <elementRef key="objectName"/>
   <elementRef key="altIdentifier"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element msIdentifier
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      (
         placeName?,
         bloc?,
         country?,
         region?,
         settlement?,
         district?,
         geogName?,
         institution?,
         repository?,
         collection*,
         idno*
      ),
      ( msName | objectName | altIdentifier )*
   )
}
```

---

**<msItem>** (manuscript item) describes an individual work or item within the intellectual content of a manuscript, manuscript part, or other object. [*11.6.1. The msItem and msItemStruct Elements*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*

    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.msClass
  - *@class*
- att.msExcerpt
  - *@defective*

*Member of* model.msItemPart

*Contained by*

*msdescription:* msContents msItem

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* author bibl biblStruct binaryObject cb cit editor ellipsis gap gb graphic index lb listBibl meeting milestone note noteGrp p pb quote respStmt textLang title

*figures:* figure notatedMusic

*header:* biblFull funder idno principal sponsor

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* colophon decoNote explicit filiation finalRubric incipit locus locusGrp msDesc msItem msItemStruct rubric

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*textstructure:* argument byline docAuthor docDate docEdition docImprint docTitle epigraph imprimatur titlePart

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Example*

```
<msItem  class="#saga">
  <locus>ff. 1r-24v</locus>
  <title>Agrip af Noregs konunga sögum</title>
  <incipit>regi oc h<ex>ann</ex> setiho
  <gap  reason="illegible"  extent="7"/>sc
    heim se<ex>m</ex> þio</incipit>
  <explicit>h<ex>on</ex> hev<ex>er</ex>
   <ex>oc</ex>þa buit hesta .ij. aNan viþ
    fé enh<ex>on</ex>o<ex>m</ex> aNan til
```

```
      reiþ<ex>ar</ex>
    </explicit>
    <textLang  mainLang="non">Old Norse/Icelandic</textLang>
  </msItem>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="locus"/>
   <elementRef key="locusGrp"/>
  </alternate>
  <alternate>
   <classRef key="model.pLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <alternate minOccurs="1"
    maxOccurs="unbounded">
    <classRef key="model.titlepagePart"/>
    <classRef key="model.msItemPart"/>
    <classRef key="model.global"/>
   </alternate>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element msItem
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.msClass.attributes,
   att.msExcerpt.attributes,
   (
      ( locus | locusGrp )*,
      (
         model.pLike+
       | ( model.titlepagePart | model.msItemPart | model.global )+
      )
   )
}
```

---

**<msItemStruct>** (structured manuscript item) contains a structured description for an individual work or item within the intellectual content of a manuscript, manuscript part, or other object. [*11.6.1. The msItem and msItemStruct Elements*]

*Module* msdescription — Manuscript Description
*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.msClass
  - *@class*
- att.msExcerpt
  - *@defective*

*Member of* model.msItemPart
*Contained by*
*msdescription:* msContents msItem msItemStruct
*May contain*
*core:* author bibl biblStruct listBibl note noteGrp p respStmt textLang title
*linking:* ab
*msdescription:* colophon decoNote explicit filiation finalRubric incipit locus locusGrp msItemStruct rubric
*Example*

```
<msItemStruct n="2" defective="false"
  class="#biblComm">
  <locus from="24v" to="97v">24v-97v</locus>
  <author>Apringius de Beja</author>
  <title type="uniform" xml:lang="la">Tractatus in Apocalypsin</title>
  <rubric>Incipit Trac<supplied reason="omitted">ta</supplied>tus
     in apoka<lb/>lipsin eruditissimi uiri <lb/> Apringi ep<ex>iscop</ex>i
     Pacensis eccl<ex>esi</ex>e</rubric>
  <finalRubric>EXPLIC<ex>IT</ex> EXPO<lb/>SITIO APOCALIPSIS
     QVA<ex>M</ex> EXPOSVIT DOM<lb/>NVS APRINGIUS EP<ex>ISCOPU</ex>S.
     DEO GR<ex>ACI</ex>AS AGO. FI<lb/>NITO LABORE ISTO.</finalRubric>
  <bibl>
    <ref target="http://amiBibl.xml#Apringius1900">Apringius</ref>, ed.
Férotin</bibl>
  <textLang mainLang="la">Latin</textLang>
</msItemStruct>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0">
   <elementRef key="locus"/>
   <elementRef key="locusGrp"/>
  </alternate>
  <alternate>
   <classRef key="model.pLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <sequence>
    <elementRef key="author" minOccurs="0"
     maxOccurs="unbounded"/>
    <elementRef key="respStmt"
     minOccurs="0" maxOccurs="unbounded"/>
    <elementRef key="title" minOccurs="0"
     maxOccurs="unbounded"/>
    <elementRef key="rubric" minOccurs="0"/>
    <elementRef key="incipit" minOccurs="0"/>
    <elementRef key="msItemStruct"
     minOccurs="0" maxOccurs="unbounded"/>
    <elementRef key="explicit"
     minOccurs="0"/>
    <elementRef key="finalRubric"
     minOccurs="0"/>
    <elementRef key="colophon"
     minOccurs="0" maxOccurs="unbounded"/>
    <elementRef key="decoNote"
     minOccurs="0" maxOccurs="unbounded"/>
    <elementRef key="listBibl"
     minOccurs="0" maxOccurs="unbounded"/>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <elementRef key="bibl"/>
     <elementRef key="biblStruct"/>
    </alternate>
    <elementRef key="filiation"
     minOccurs="0"/>
    <classRef key="model.noteLike"
     minOccurs="0" maxOccurs="unbounded"/>
```

```
    <elementRef key="textLang"
     minOccurs="0"/>
   </sequence>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element msItemStruct
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.msClass.attributes,
   att.msExcerpt.attributes,
   (
      ( locus | locusGrp )?,
      (
         model.pLike+
       | (
            author*,
            respStmt*,
            title*,
            rubric?,
            incipit?,
            msItemStruct*,
            explicit?,
            finalRubric?,
            colophon*,
            decoNote*,
            listBibl*,
            ( bibl | biblStruct )*,
            filiation?,
            model.noteLike*,
            textLang?
         )
      )
   )
}
```

---

## <msName>

(manuscript name) contains a proper noun or noun phrase used for a manuscript, or other object, as opposed to a formal identification number or classmark. [*11.4. The Manuscript Identifier*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.typed
  - *@type*
  - *@subtype*

*Contained by:* *msDescription:* msIdentifier
*namesdates:* objectIdentifier
*May contain*
*core:* name rs
*gaiji:* g
character data
*Example*

```
<msName>The Vercelli Book</msName>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
```

```
  <textNode/>
  <classRef key="model.gLike"/>
  <elementRef key="rs"/>
  <elementRef key="name"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element msName
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   ( text | model.gLike | rs | name )*
}
```

---

**<msPart>** (manuscript part) contains information about an originally distinct manuscript or part of a manuscript, which is now part of a composite manuscript. [*11.10. Manuscript Parts*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.typed
  - *@type*
  - *@subtype*

*Contained by:* *msDescription:* msDesc msPart

*May contain*

*core:* head p

*linking:* ab

*msdescription:* additional history msContents msIdentifier msPart physDesc

*Note* As this last example shows, for compatibility reasons the identifier of a manuscript part may be supplied as a simple <altIdentifier> rather than using the more structured <msIdentifier> element. This usage is however deprecated.

*Example*

```
<msPart>
  <msIdentifier>
    <idno>A</idno>
    <altIdentifier type="catalog">
      <collection>Becker</collection>
      <idno>48, Nr. 145</idno>
    </altIdentifier>
    <altIdentifier type="catalog">
      <collection>Wiener Liste</collection>
      <idno>4°5</idno>
    </altIdentifier>
  </msIdentifier>
  <head>
    <title xml:lang="la">Gregorius: Homiliae in Ezechielem</title>
    <origPlace key="tgn_7008085">Weissenburg (?)</origPlace>
    <origDate notBefore="0801"
      notAfter="0815">IX. Jh., Anfang</origDate>
  </head>
</msPart>
```

*Example*

```
<msDesc>
  <msIdentifier>
    <settlement>Amiens</settlement>
    <repository>Bibliothèque Municipale</repository>
    <idno>MS 3</idno>
```

```
      <msName>Maurdramnus Bible</msName>
    </msIdentifier>
    <msContents>
      <summary>
        <list>
          <item xml:lang="en">Miscellany of various texts;</item>
          <item xml:lang="la">Prudentius, Psychomachia;</item>
          <item xml:lang="la">Physiologus de natura animantium</item>
        </list>
      </summary>
      <textLang mainLang="la">Latin</textLang>
    </msContents>
    <physDesc>
      <objectDesc form="composite_manuscript"/>
    </physDesc>
    <msPart>
      <msIdentifier>
        <idno>ms. 10066-77 ff. 140r-156v</idno>
      </msIdentifier>
      <msContents>
        <summary xml:lang="la">Physiologus</summary>
        <textLang mainLang="la">Latin</textLang>
      </msContents>
    </msPart>
    <msPart>
      <msIdentifier>
        <altIdentifier>
          <idno>MS 6</idno>
        </altIdentifier>
      </msIdentifier>
  <!-- other information specific to this part here -->
    </msPart>
  <!-- more parts here -->
  </msDesc>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="msIdentifier"/>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate>
   <classRef key="model.pLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="msContents"/>
    <elementRef key="physDesc"/>
    <elementRef key="history"/>
    <elementRef key="additional"/>
    <elementRef key="msPart"/>
   </alternate>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element msPart
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   (
      msIdentifier,
      model.headLike*,
      (
         model.pLike+
       | ( msContents | physDesc | history | additional | msPart )*
      )
   )
}
```

---

## <musicNotation>

(music notation) contains description of type of musical notation. [*11.7.2. Writing, Decoration, and Other Notations*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.physDescPart

*Contained by*

*msdescription:* physDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<musicNotation>
  <p>Square notation of 4-line red staves.</p>
</musicNotation>
```

*Example*

```
<musicNotation>Neumes in <term>campo aperto</term> of the St. Gall type.
</musicNotation>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element musicNotation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.specialPara}
```

---

## <name>

(name, proper noun) contains a proper noun or noun phrase. [*3.6.1. Referring Strings*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*

    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.personal
  - *@full*
  - *@sort*

- att.naming
  - *@role*
  - *@nymRef*
- att.canonical
  - *@key*
  - *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.nameLike.agent model.personPart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp respStmt rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose setting

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msName musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona place placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap

gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Proper nouns referring to people, places, and organizations may be tagged instead with <persName>, <placeName>, or <orgName>, when the TEI module for names and dates is included.

*Example*

```
<name type="person">Thomas Hoccleve</name>
<name type="place">Villingaholt</name>
<name type="org">Vetus Latina Institut</name>
<name type="person" ref="#HOC001">Occleve</name>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element name
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
```

```
   att.editLike.attributes,
   att.personal.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<nameLink>** (name link) contains a connecting phrase or link used within a name but not regarded as part of it, such as *van der* or *of*. [*14.2.1. Personal Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*

- att.typed
  - *@type*
  - *@subtype*

*Member of* model.persNamePart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<persName>
  <forename>Frederick</forename>
  <nameLink>van der</nameLink>
  <surname>Tronck</surname>
</persName>
```

*Example*

```
<persName>
  <forename>Alfred</forename>
  <nameLink>de</nameLink>
  <surname>Musset</surname>
</persName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element nameLink
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<namespace>** (namespace) supplies the formal name of the namespace to which the elements documented by its children belong. [*2.3.4. The Tagging Declaration*]

*Module* header — The TEI Header
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  * *@ana*
- att.global.change
  * *@change*
- att.global.facs
  * *@facs*
- att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

`@name` specifies the full formal name of the namespace concerned.
*Status* Required
*Datatype* 0–1 occurrences of teidata.namespace separated by whitespace

*Contained by:* tagsDecl
*May contain*
*header:* tagUsage
*Example*

```
<namespace name="http://www.tei-c.org/ns/1.0">
  <tagUsage gi="hi" occurs="28" withId="2"> Used only to mark English words
```

```
        italicized in the copy text </tagUsage>
    </namespace>
```

*Content model*

```
<content>
 <elementRef key="tagUsage" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element namespace
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute name { ? },
   tagUsage+
}
```

---

## <nationality>

<nationality> (nationality) contains an informal description of a person's present or past nationality or citizenship. [*16.2.2. The Participant Description*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*

    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*

- att.canonical
  - *@key*
  - *@ref*
- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **birth**
**naturalised**
**self-assigned**

*Member of* model.persStateLike
*Contained by*
*namesdates:* person personGrp persona
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit
*dictionaries:* lang oRef pRef
*figures:* figure formula notatedMusic
*gaiji:* g
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef specDesc specList tag val
*textcrit:* app witDetail
*textstructure:* floatingText
*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo
*verse:* caesura rhyme
character data

*Example*

```
<nationality key="US" notBefore="1966"> Obtained US Citizenship in
1966</nationality>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element nationality
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.phraseSeq}
```

---

**<node>** (node) encodes a node, a possibly labeled point in a graph. [*20.1. Graphs and Digraphs*]

*Module* nets — Graphs, Networks, and Trees

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*

  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

- att.typed
  - ~~type~~
  - @subtype

`@type` provides a type for a node.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **initial** initial node in a transition network

**final** final node in a transition network

```
<graph>
  <node  xml:id="gnex11"
    adjTo="#gnex12"
    adj="#gnex13"
    type="initial">
    <label>1</label>
  </node>
  <node  xml:id="gnex12"
    adjFrom="#gnex11"
    adjTo="#gnex13">
    <label>2</label>
  </node>
  <node  xml:id="gnex13"
    adjFrom="#gnex12"
    adj="#gnex11"
    type="final">
    <label>3</label>
  </node>
</graph>
```

`@value` provides the value of a node, which is a feature structure or other analytic element.

*Status* Optional

*Datatype* teidata.pointer

`@adjTo` (adjacent to) gives the identifiers of the nodes which are adjacent to the current node.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

```
<graph>
  <node  xml:id="gnex21"
    adjTo="#gnex22"
```

```
      adj="#gnex23"
      type="initial">
      <label>1</label>
    </node>
    <node  xml:id="gnex22"
      adjTo="#gnex23">
      <label>2</label>
    </node>
    <node  xml:id="gnex23"
      type="final">
      <label>3</label>
    </node>
  </graph>
```

`@adjFrom` (adjacent from) gives the identifiers of the nodes which are adjacent from the current node.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

```
  <graph>
    <node  xml:id="gnex31"
      type="initial">
      <label>1</label>
    </node>
    <node  xml:id="gnex32"
      adjFrom="#gnex31">
      <label>2</label>
    </node>
    <node  xml:id="gnex33"
      adjFrom="#gnex32"
      adj="#gnex31"
      type="final">
      <label>3</label>
    </node>
  </graph>
```

`@adj` (adjacent) gives the identifiers of the nodes which are both adjacent to and adjacent from the current node.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

```
  <graph>
    <node  xml:id="gnex41"
      adj="#gnex42 #gnex43"
      type="initial">
      <label>1</label>
    </node>
    <node  xml:id="gnex42"
      adj="#gnex41 #gnex43">
      <label>2</label>
    </node>
    <node  xml:id="gnex43"
      adj="#gnex42 #gnex41"
      type="final">
      <label>3</label>
    </node>
  </graph>
```

*Note* Use this attribute instead of the *adjTo* and *adjFrom* attributes when the graph is undirected and vice versa if the graph is directed.

`@inDegree` gives the in degree of the node, the number of nodes which are adjacent from the given node.

*Status* Optional

*Datatype* teidata.count

```
<graph>
  <node  xml:id="gnex51"
    adjTo="#gnex52"
    adj="#gnex53"
    inDegree="1"
    type="initial">
    <label>1</label>
  </node>
  <node  xml:id="gnex52"
    adjFrom="#gnex51"
    adjTo="#gnex53"
    inDegree="2">
    <label>2</label>
  </node>
  <node  xml:id="gnex53"
    adjFrom="#gnex52"
    adj="#gnex51"
    inDegree="2"
    type="final">
    <label>3</label>
  </node>
</graph>
```

`@outDegree` gives the out degree of the node, the number of nodes which are adjacent to the given node.

*Status* Optional

*Datatype* teidata.count

```
<graph>
  <node  xml:id="gnex61"
    adjTo="#gnex62"
    adj="#gnex63"
    outDegree="2"
    type="initial">
    <label>1</label>
  </node>
  <node  xml:id="gnex62"
    adjFrom="#gnex61"
    adjTo="#gnex63"
    outDegree="1">
    <label>2</label>
  </node>
  <node  xml:id="gnex63"
    adjFrom="#gnex62"
    adj="#gnex61"
    outDegree="2"
    type="final">
    <label>3</label>
  </node>
</graph>
```

`@degree` gives the degree of the node, the number of arcs with which the node is incident.

*Status* Optional

*Datatype* teidata.count

```
<graph>
```

```
      <node  xml:id="gnex71"
        adjTo="#gnex72"
        adj="#gnex73"
        outDegree="1"
        inDegree="0"
        degree="1"
        type="initial">
        <label>1</label>
      </node>
      <node  xml:id="gnex72"
        adjFrom="#gnex71"
        adjTo="#gnex73"
        outDegree="1"
        inDegree="1"
        degree="0">
        <label>2</label>
      </node>
      <node  xml:id="gnex73"
        adjFrom="#gnex72"
        adj="#gnex71"
        outDegree="0"
        inDegree="1"
        degree="1"
        type="final">
        <label>3</label>
      </node>
    </graph>
```

*Note* Use this attribute instead of the *inDegree* and *outDegree* attributes when the graph is undirected and vice versa if the graph is directed.

*Contained by* graph

*May contain*

*core:* label

*Note* Zero, one, or two children <label> elements may be present. The first occurrence of <label> provides a label for the arc; the second provides a second label for the arc, and should be used if a transducer is being encoded whose actions are associated with nodes rather than with arcs.

*Example*

```
<node  xml:id="t6"  type="final"  inDegree="2"
  outDegree="0">
  <label>6</label>
</node>
```

*Content model*

```
<content>
 <sequence minOccurs="0">
  <elementRef key="label"/>
  <elementRef key="label" minOccurs="0"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element node
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { "initial" | "final" }?,
   attribute value { text }?,
   attribute adjTo { list { + } }?,
   attribute adjFrom { list { + } }?,
   attribute adj { list { + } }?,
   attribute inDegree { text }?,
   attribute outDegree { text }?,
   attribute degree { text }?,
   ( ( label, label? )? )
}
```

---

**<normalization>** (normalization) indicates the extent of normalization or regularization of the original source carried out in converting it to electronic form. [*2.3.3. The Editorial Practices Declaration 16.3.2. Declarable Elements*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

  * *@select*
  - att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
  - att.global.responsibility
  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.declarable
  - *@default*

`@method` indicates the method adopted to indicate normalizations within the text.
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **silent** normalization made silently[Default]
**markup** normalization represented using markup

*Member of* model.editorialDeclPart
*Contained by*
*header:* editorialDecl
*May contain*
*core:* p
*linking:* ab
*Example*

```
<editorialDecl>
  <normalization  method="markup">
    <p>Where both upper- and lower-case i, j, u, v, and vv have been
normalized, to modern
        20th century typographical practice, the <gi>choice</gi> element has
been used to
        enclose <gi>orig</gi> and <gi>reg</gi> elements giving the original and
new values
        respectively. ... </p>
  </normalization>
  <normalization  method="silent">
    <p>Spacing between words and following punctuation has been regularized to
zero spaces;
        spacing between words has been regularized to one space.</p>
  </normalization>
  <normalization  source="http://www.dict.sztaki.hu/webster">
    <p>Spelling converted throughout to Modern American usage, based on
Websters 9th
        Collegiate dictionary.</p>
  </normalization>
</editorialDecl>
```

*Schematron* <sch:pattern is-a=”declarable”> <sch:param name=”tde” value=”tei:normalization”/>
</sch:pattern>
*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element normalization
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   attribute method { "silent" | "markup" }?,
   model.pLike+
}
```

---

**<notatedMusic>** encodes the presence of music notation in a text. [*15.3. Notated Music in Written Text*]

*Module* figures — Tables, Formulæ, Graphics, and Notated Music

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.placement
  - *@place*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.global

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor

docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*core:* binaryObject desc graphic label ptr ref

*linking:* seg

*tagdocs:* listRef

*Note* It is possible to describe the content of the notation using elements from the `model.labelLike` class and it is possible to point to an external representation using elements from `model.ptrLike`. It is possible to specify the location of digital objects representing the notated music in other media such as images or audio-visual files. The encoder's interpretation of the correspondence between the notated music and these digital objects is not encoded explicitly. We recommend the use of graphic and binaryObject mainly as a fallback mechanism when the notated music format is not displayable by the application using the encoding. The alignment of encoded notated music, images carrying the notation, and audio files is a complex matter for which we refer the encoder to other formats and specifications such as MPEG-SMR.

It is also recommended, when useful, to embed XML-based music notation formats, such as the Music Encoding Initiative format as content of <notatedMusic>. This must be done by means of customization.

*Example*

```
<notatedMusic>
  <ptr  target="bar1.xml"/>
  <graphic  url="bar1.jpg"/>
  <desc>First bar of Chopin's Scherzo No.3 Op.39</desc>
</notatedMusic>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.labelLike"/>
  <classRef key="model.ptrLike"/>
  <elementRef key="graphic"/>
  <elementRef key="binaryObject"/>
  <elementRef key="seg"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element notatedMusic
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.placement.attributes,
```

```
   att.typed.attributes,
   ( model.labelLike | model.ptrLike | graphic | binaryObject | seg )*
}
```

---

**<note>** (note) contains a note or annotation. [*3.9.1. Notes and Simple Annotation 2.2.6. The Notes Statement 3.12.2.8. Notes and Statement of Language 10.3.5.4. Notes within Entries*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.anchoring
  - *@anchored*
  - *@targetEnd*
- att.cmc

  - *@generatedBy*
- att.placement
  - *@place*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*

*Member of* model.annotationLike model.annotationPart.body model.correspActionPart model.correspContextPart model.correspDescPart model.noteLike

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope biblStruct cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned monogr name note noteGrp num orig p pubPlace publisher q quote rb ref reg resp respStmt rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*gaiji:* char glyph

*header:* authority change classCode correspAction correspContext correspDesc distributor edition extent funder geoDecl handNote language licence notesStmt principal scriptNote sponsor

*linking:* ab annotation listAnnotation seg standOff

*msdescription:* accMat acquisition additions adminInfo altIdentifier catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem msItemStruct musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc climate country death district education event eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality object objectName occupation offset org orgName persName persPronouns person personGrp persona place placeName population region residence roleName settlement sex socecStatus state surname terrain trait

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* app lem rdg rdgGrp wit witDetail witness

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metDecl metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example* In the following example, the translator has supplied a footnote containing an explanation of the term translated as ”painterly”:

```
And yet it is not only
in the great line of Italian renaissance art, but even in the
painterly <note  place="bottom"  type="gloss"
  resp="#MDMH">
  <term  xml:lang="de">Malerisch</term>. This word has, in the German, two
distinct meanings, one objective, a quality residing in the object,
the other subjective, a mode of apprehension and creation. To avoid
confusion, they have been distinguished in English as
<mentioned>picturesque</mentioned> and
<mentioned>painterly</mentioned> respectively.
</note> style of the
Dutch genre painters of the seventeenth century that drapery has this
psychological significance.

<!-- elsewhere in the document -->
<respStmt  xml:id="MDMH">
  <resp>translation from German to English</resp>
```

```
  <name>Hottinger, Marie Donald Mackie</name>
</respStmt>
```

For this example to be valid, the code MDMH must be defined elsewhere, for example by means of a responsibility statement in the associated TEI header.

*Example* The global *n* attribute may be used to supply the symbol or number used to mark the note's point of attachment in the source text, as in the following example:

```
Mevorakh b. Saadya's mother, the matriarch of the
family during the second half of the eleventh century,
<note  n="126"  anchored="true"> The
alleged mention of Judah Nagid's mother in a letter from 1071 is, in fact, a
reference to
Judah's children; cf. above, nn. 111 and 54. </note> is well known from Geniza
documents
published by Jacob Mann.
```

However, if notes are numbered in sequence and their numbering can be reconstructed automatically by processing software, it may well be considered unnecessary to record the note numbers.

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element note
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.anchoring.attributes,
   att.cmc.attributes,
   att.placement.attributes,
   att.pointing.attributes,
   att.typed.attributes,
   att.written.attributes,
   macro.specialPara}
```

---

## <noteGrp>

<noteGrp> (note group) contains a group of notes. [*3.9.1.1. Encoding Grouped Notes*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.anchoring
  - *@anchored*
  - *@targetEnd*
- att.cmc
  - *@generatedBy*
- att.placement
  - *@place*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*

*Member of* model.correspActionPart model.correspContextPart model.correspDescPart model.noteLike

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope biblStruct cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned monogr name note noteGrp num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*gaiji:* char glyph

*header:* authority change classCode correspAction correspContext correspDesc distributor edition extent funder geoDecl handNote language licence notesStmt principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions adminInfo catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem msItemStruct musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc climate country death district education event eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality object objectName occupation offset org orgName persName persPronouns person personGrp persona place placeName population region residence roleName settlement sex socecStatus state surname terrain trait

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* app lem rdg rdgGrp wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metDecl metSym rhyme

*May contain*

*core:* desc note noteGrp

*Example* In the following example, there are two notes in different languages, each specifying the content of the annotation relating to the same fragment of text:

```
<p>(...) tamen reuerendos dominos archiepiscopum et canonicos Leopolienses
necnon episcopum in duplicibus Quatuortemporibus
<noteGrp>
   <note  xml:lang="en">Quatuor Tempora, so called dry fast days (Wednesday,
Friday, and Saturday)
        falling on each of the quarters of the year. In the first quarter they
were called Cinerum
        (following Ash Wednesday), second Spiritus (following Pentecost), third
Crucis
```

```
        (after the Exaltation of the Holy Cross, September 14th), and Luciae
        in the fourth (after the feast of St. Lucia, December 13th).
   </note>
   <note  xml:lang="pl">Quatuor Tempora, tzw. Suche dni postne (środa, piątek i
sobota)
        przypadające cztery razy w roku. W pierwszym kwartale zwały się Cinerum
        (po Popielcu), w drugim Spiritus (po Zielonych Świętach), w trzecim
Crucis
        (po święcie Podwyższenia Krzyża 14 września), w czwartym Luciae
        (po dniu św. Łucji 13 grudnia).
   </note>
  </noteGrp>
totaliter expediui.
</p>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="desc" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="note"/>
   <elementRef key="noteGrp"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element noteGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.anchoring.attributes,
   att.cmc.attributes,
   att.placement.attributes,
   att.pointing.attributes,
   att.typed.attributes,
   att.written.attributes,
   ( desc*, ( note | noteGrp )+ )
}
```

---

**<notesStmt>** (notes statement) collects together any notes providing information about a text additional to that recorded in other parts of the bibliographic description. [*2.2.6. The Notes Statement 2.2. The File Description*]

*Module* header — The TEI Header
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  * *@ana*
- att.global.change
  * *@change*
- att.global.facs
  * *@facs*
- att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

*Contained by* *header:* biblFull fileDesc
*May contain*
*core:* note noteGrp relatedItem
*Note* Information of different kinds should not be grouped together into the same note.
*Example*

```
<notesStmt>
  <note>Historical commentary provided by Mark Cohen</note>
  <note>OCR scanning done at University of Toronto</note>
</notesStmt>
```

*Content model*

```
<content>
```

```
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.noteLike"/>
  <elementRef key="relatedItem"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element notesStmt
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.noteLike | relatedItem )+
}
```

---

**<num>** (number) contains a number, written in any form. [*3.6.3. Numbers and Measures*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.ranging
  - *@atLeast*
  - *@atMost*
  - *@min*
  - *@max*
  - *@confidence*
- att.typed
  - ~~type~~
  - @subtype

`@type` indicates the type of numeric value.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **cardinal** absolute number, e.g. 21, 21.5

**ordinal** ordinal number, e.g. 21st

**fraction** fraction, e.g. one half or three-quarters

**percentage** a percentage

*Note* If a different typology is desired, other values can be used for this attribute.

`@value` supplies the value of the number in standard form.

*Status* Optional

*Datatype* teidata.numeric

*Values* a numeric value.

*Note* The standard form used is defined by the TEI datatype `teidata.numeric`.

*Member of* model.measureLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure measureGrp meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Detailed analyses of quantities and units of measure in historical documents may also use the feature

structure mechanism described in chapter *19. Feature Structures.* The <num> element is intended for use in simple applications.

*Example*

```
<p>I reached <num  type="cardinal"  value="21">twenty-one</num> on
my <num  type="ordinal"  value="21">twenty-first</num> birthday</p>
<p>Light travels at <num  value="3E10">3×10<hi  rend="sup">10</hi>
  </num> cm per second.</p>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element num
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.ranging.attributes,
   att.typed.attribute.subtype,
   attribute type { "cardinal" | "ordinal" | "fraction" | "percentage" }?,
   attribute value { text }?,
   macro.phraseSeq}
```

---

## <number>

(number) indicates grammatical number associated with a form, as given in a dictionary. [*10.3.1. Information on Written and Spoken Forms 10.3.2. Grammatical Information*]

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*

- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*
- att.lexicographic.normalized
  - *@norm*
  - *@orig*

*Member of* model.morphLike

*Contained by*

*dictionaries:* dictScrap entryFree form gramGrp

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* This element is synonymous with <gram type="num">.

*Example*

```
<entry>
  <form>
    <orth>wits</orth>
    <pron>wIts</pron>
  </form>
  <gramGrp>
    <number>pl</number>
    <pos>n</pos>
  </gramGrp>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element number
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
```

```
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   macro.paraContent}
```

---

**<numeric>** (numeric value) represents the value part of a feature-value specification which contains a numeric value or range. [*19.3. Other Atomic Feature Values*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*

`@value` supplies a lower bound for the numeric value represented, and also (if *max* is not supplied) its upper bound.
*Status* Required
*Datatype* teidata.numeric

`@max` supplies an upper bound for the numeric value represented.
*Status* Optional
*Datatype* teidata.numeric

`@trunc` specifies whether the value represented should be truncated to give an integer value.
*Status* Optional
*Datatype* teidata.truthValue

*Member of* model.featureVal.single
*Contained by*
*iso-fs:* f fvLib if vAlt vColl vDefault vLabel vMerge vNot vRange
*May contain* Empty element
*Note* It is an error to supply the *max* attribute in the absence of a value for the *value* attribute.
*Example*

```
<numeric value="42"/>
```

This represents the numeric value 42.
*Example*

```
<numeric value="42.45" max="50" trunc="true"/>
```

This represents any of the nine possible integer values between 42 and 50 inclusive. If the *trunc* attribute had the value FALSE, this example would represent any of the infinite number of numeric values between 42.45 and 50.0
*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element numeric
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datcat.attributes,
   attribute value { text },
   attribute max { text }?,
   attribute trunc { text }?,
   empty
}
```

---

## <nym>

(canonical name) contains the definition for a canonical name or name component of any kind. [*14.3.7. Names and Nyms*]

*Module* namesdates — Names, Dates, People, and Places
*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

`@parts` points to constituent nyms.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Contained by:* *andnysdates:* listNym nym

*May contain*

*core:* p

*dictionaries:* colloc def etym form gramGrp hom hyph lbl orth pos pron re sense subc superEntry syll usg xr

*header:* idno
*linking:* ab
*namesdates:* nym
*Example*

```
<nym  xml:id="J452">
  <form>
    <orth  xml:lang="en-US">Ian</orth>
    <orth  xml:lang="en-x-Scots">Iain</orth>
  </form>
</nym>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="idno" minOccurs="0"
   maxOccurs="unbounded"/>
  <classRef key="model.entryPart"
   minOccurs="0" maxOccurs="unbounded"/>
  <classRef key="model.pLike" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="nym" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element nym
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   attribute parts { list { + } }?,
   ( idno*, model.entryPart*, model.pLike*, nym* )
}
```

---

## <object> contains a description of a single identifiable physical object. [*14.3.6. Objects*]

*Module* namesdates — Names, Dates, People, and Places
*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.canonical
  - *@key*
  - *@ref*
- att.declaring
  - *@decls*
- att.docStatus
  - *@status*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.objectLike
*Contained by*
*namesdates:* event listObject object
*textcrit:* witness
*May contain*

*core:* bibl biblStruct head listBibl note noteGrp p

*header:* biblFull

*linking:* ab link linkGrp

*msdescription:* additional history msContents msDesc physDesc

*namesdates:* object objectIdentifier

*Note* The <object> element is a recent addition to the TEI P5 Guidelines as of version 3.5.0 and as such may be more prone to further revision in the next few releases as its use develops. This may be particularly evident where its contents have been borrowed from <msDesc> and have yet to be generalized from their use in the context of manuscript descriptions.

The <object> element may be used for describing any object, text-bearing or not, though where the textuality of the object is the primary concern or a collection is mostly composed of manuscripts, encoders may prefer the <msDesc> element (a more specific form of <object>) which may be used not only to describe manuscripts but any form of text-bearing objects such as early printed books.

*Example*

```
<listObject>
  <object  xml:id="Alfred_Jewel">
    <objectIdentifier>
      <country>United Kingdom</country>
      <region>Oxfordshire</region>
      <settlement>Oxford</settlement>
      <institution>University of Oxford</institution>
      <repository>Ashmolean Museum</repository>
      <collection>English Treasures</collection>
      <idno  type="ashmolean">AN1836p.135.371</idno>
      <idno  type="wikipedia">https://en.wikipedia.org/wiki/Alfred_Jewel</idno>
      <objectName>Alfred Jewel</objectName>
    </objectIdentifier>
    <physDesc>
      <p> The Alfred Jewel is about 6.4 cm in length and is made of combination
of filigreed <material>gold</material>
            surrounding a polished teardrop shaped piece of transparent
<material>quartz</material>. Underneath the rock
            crystal is a cloisonné enamel image of a man with ecclesiastical
symbols. The sides of the jewel holding the
            crystal in place contain an openwork inscription saying "AELFRED MEC
HEHT GEWYRCAN", meaning 'Alfred ordered
            me made'. </p>
    </physDesc>
    <history>
      <origin>It is generally accepted that the Alfred Jewel dates from the
<origDate>late 9th Century</origDate> and
            was most likely made in <origPlace>England</origPlace>. </origin>
      <provenance  when="1693">The jewel was discovered in 1693 at Petherton
Park, North Petherton in the English
            county of Somerset, on land owned by Sir Thomas Wroth. North
Petherton is about 8 miles away from Athelney,
            where King Alfred founded a monastery. </provenance>
      <provenance  when="1698">A description of the Alfred Jewel was first
published in 1698, in the Philosophical
            Transactions of the Royal Society.</provenance>
      <acquisition> It was bequeathed to Oxford University by Colonel Nathaniel
Palmer (c. 1661-1718) and today is in
            the Ashmolean Museum in Oxford. </acquisition>
    </history>
  </object>
```

```
</listObject>
<!-- Elsewhere in document -->
<p> The <objectName  ref="#MinsterLovellJewel">Minster Lovell
Jewel</objectName> is probably the most similar to the
<objectName  ref="#Alfred_Jewel">Alfred Jewel</objectName> and was found in
<placeName  ref="#MinsterLovell">Minster
    Lovell</placeName> in
<placeName  ref="#Oxfordshire">Oxfordshire</placeName> and is kept at the
<orgName  ref="#AshmoleanMuseum">Ashmolean Museum</orgName>.
</p>
```

*Example*

```
<listObject>
  <object  xml:id="MaskOfTutankhamun">
    <objectIdentifier>
      <objectName  xml:lang="en">Mask of Tutankhamun</objectName>
      <idno  type="carter">256a</idno>
      <idno  type="JournalD'Entrée">60672</idno>
      <idno  type="exhibition">220</idno>
      <institution>Museum of Egyptian Antiquities</institution>
      <address>
        <street>15 Meret Basha</street>
        <district>Ismailia</district>
        <settlement>Cairo</settlement>
        <country>Egypt</country>
        <location>
          <geo>30.047778 31.233333</geo>
        </location>
      </address>
    </objectIdentifier>
    <msContents>
      <p>The back and shoulders of the mask is inscribed with a protective spell
in Egyptian hieroglyphs formed of ten
           vertical and horizontal lines. This spell first appeared on masks in
the Middle Kingdom at least 500 years
           before Tutankhamun, and comes from chapter 151 of the <title>Book of
the Dead</title>.</p>
    </msContents>
    <physDesc>
      <p> The mask of Tutankhamun is 54cm x 39.3cm x 49cm. It is constructed
from two layers of high-karat gold that
           varies in thickness from 1.5-3mm. It weighs approximately 10.23kg
and x-ray crystallography shows that it is
           composed of two alloys of gold with a lighter 18.4 karat shade being
used for the face and neck while a heavier
           22.5 karat gold was used for the rest of the mask.</p>
      <p>In the mask Tutankhamun wears a nemes headcloth which has the royal
insignia of a cobra (Wadjet) and vulture
           (Nekhbet) on it. These are thought respectively to symbolize
Tutankhamun's rule of both Lower Egypt and Upper
           Egypt. His ears are pierced for earrings. The mask has rich inlays
of coloured glass and gemstones, including
           lapis lazuli surrounding the eye and eyebrows, quartz for the eyes,
obsidian for the pupils. The broad collar is
           made up of carnelian, feldspar, turquoise, amazonite, faience and
other stones.</p>
```

```
    </physDesc>
    <history>
      <origin>
        <p>The mask of Tutankhamun was created in <origPlace>Egypt</origPlace>
around <origDate  when="-1323"  type="circa">1323 BC</origDate>. It is a death
mask of the 18th-dynasty ancient Egyptian Pharaoh Tutankhamun
                who reigned 1332–1323 BC. </p>
      </origin>
      <provenance>
        <p>The mask of Tutankhamun was found in his burial chamber at Theban
Necropolis in the Valley of the Kings in
                1922. On 28 October 1925 the excavation team led by English
archaeologist Howard Carter opened the heavy
                sarcophagus and three coffins and were the first people in around
3,250 years to see the mask of Tutankhamun.
                Carter wrote in his diary: <quote> The pins removed, the lid was
raised. The penultimate scene was disclosed —
                    a very neatly wrapped mummy of the young king, with golden
mask of sad but tranquil expression, symbolizing
                    Osiris … the mask bears that god's attributes, but the
likeness is that of Tut.Ankh.Amen — placid and
                    beautiful, with the same features as we find upon his statues
and coffins. The mask has fallen slightly
                    back, thus its gaze is straight up to the heavens. </quote>
        </p>
      </provenance>
      <acquisition> In December 1925, the mask was removed from the tomb, placed
in a crate and transported 635
            kilometres (395 mi) to the Egyptian Museum in Cairo, where it
remains on public display. </acquisition>
    </history>
    <additional>
      <adminInfo>
        <custodialHist>
          <custEvent  when="1944">When it was discovered in 1925, the 2.5kg
narrow gold beard was no longer attached to
                    the mask and was reattached to the chin by use of a wooden
dowel in 1944.</custEvent>
          <custEvent  when="2014-08"> In August 2014 when the mask was removed
from its display case for cleaning, the
                    beard fell off again. Those working in the museum unadvisedly
used a quick-drying epoxy to attempt to fix
                    it, but left the beard off-centre. </custEvent>
          <custEvent  when="2015-01">The damage was noticed and repaired in
January 2015 by a German-Egyptian team who
                    used beeswax, a material known to be used as adhesives by the
ancient Egyptians.</custEvent>
        </custodialHist>
      </adminInfo>
    </additional>
  </object>
</listObject>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="objectIdentifier"
```

```
   minOccurs="1" maxOccurs="unbounded"/>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate>
   <classRef key="model.pLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="msContents"/>
    <elementRef key="physDesc"/>
    <elementRef key="history"/>
    <elementRef key="additional"/>
   </alternate>
  </alternate>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.noteLike"/>
   <classRef key="model.biblLike"/>
   <elementRef key="linkGrp"/>
   <elementRef key="link"/>
  </alternate>
  <elementRef key="object" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element object
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.declaring.attributes,
   att.docStatus.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      objectIdentifier+,
      model.headLike*,
      ( model.pLike+ | ( msContents | physDesc | history | additional )* ),
      ( model.noteLike | model.biblLike | linkGrp | link )*,
      object*
   )
}
```

---

**<objectDesc>** (object description) contains a description of the physical components making up the object which is being described. [*11.7.1. Object Description*]

*Module* msdescription — Manuscript Description
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  * *@ana*
- att.global.change
  * *@change*
- att.global.facs
  * *@facs*
- att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

`@form` (form) a short project-specific name identifying the physical form of the carrier, for example as a codex, roll, fragment, partial leaf, cutting etc.
*Status* Optional
*Datatype* teidata.enumerated
*Note* Definitions for the terms used may typically be provided by a <valList> element in the project schema specification.

*Member of* model.physDescPart
*Contained by*
*msdescription:* physDesc
*May contain*
*core:* p
*linking:* ab

*msdescription:* layoutDesc supportDesc
*Example*

```
<objectDesc  form="codex">
  <supportDesc  material="mixed">
    <p>Early modern
    <material>parchment</material> and
    <material>paper</material>.</p>
  </supportDesc>
  <layoutDesc>
    <layout  ruledLines="25 32"/>
  </layoutDesc>
</objectDesc>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="supportDesc"
    minOccurs="0"/>
   <elementRef key="layoutDesc"
    minOccurs="0"/>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element objectDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute form { text }?,
   ( model.pLike+ | ( supportDesc?, layoutDesc? ) )
}
```

---

## <objectIdentifier> (object identifier) groups one or more identifiers or pieces of locating information concerning a single object. [*14.3.6. Objects*]

*Module* namesdates — Names, Dates, People, and Places
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.biblPart
*Contained by*
*core:* bibl
*namesdates:* object
*May contain*
*core:* address
*header:* idno
*msdescription:* altIdentifier collection institution msName repository
*namesdates:* bloc country district geogName objectName placeName region settlement
*Example*

```
<objectIdentifier>
  <country>United Kingdom</country>
  <region>Oxfordshire</region>
  <settlement>Oxford</settlement>
  <institution>University of Oxford</institution>
  <repository>Ashmolean Museum</repository>
  <collection>English Treasures</collection>
  <idno  type="ashmolean">AN1836p.135.371</idno>
  <idno  type="wikipedia">https://en.wikipedia.org/wiki/Alfred_Jewel</idno>
```

```
      <objectName>Alfred Jewel</objectName>
    </objectIdentifier>
```

*Example*

```
<object  xml:id="Excalibur-MultipleNames">
  <objectIdentifier>
    <objectName  type="main">Excalibur</objectName>
    <objectName  type="alt">Caliburn</objectName>
    <objectName  xml:lang="cy">Caledfwlch</objectName>
    <objectName  xml:lang="cnx">Calesvol</objectName>
    <objectName  xml:lang="br">Kaledvoulc'h</objectName>
    <objectName  xml:lang="la">Caliburnus</objectName>
    <country>Wales</country>
  </objectIdentifier>
  <p>Excalibur is the name for the legendary sword of King Arthur, in Welsh it
is called Caledfwlch,
    in Cornish it is called Calesvol, in Breton it is called Kaledvoulc'h, and
in Latin it is called Caliburnus.
    In some versions Excalibur's blade was engraved with phrases on opposite
sides which in translation read:
    "Take me up" and "Cast me away" (or similar).</p>
</object>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.placeNamePart"/>
  <elementRef key="institution"/>
  <elementRef key="repository"/>
  <elementRef key="collection"/>
  <elementRef key="idno"/>
  <elementRef key="msName"/>
  <elementRef key="objectName"/>
  <elementRef key="altIdentifier"/>
  <elementRef key="address"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element objectIdentifier
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      model.placeNamePart    | institution    | repository    | collection    | idno    |
}
```

---

**<objectName>** (name of an object) contains a proper noun or noun phrase used to refer to an object. [*14.2.5. Object Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*

    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.personal
  - *@full*
  - *@sort*
  - att.naming
    * *@role*
    * *@nymRef*
    * att.canonical
      · *@key*
      · *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.nameLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msIdentifier musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectIdentifier objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<p>
The <objectName  ref="#MinsterLovellJewel">Minster Lovell Jewel</objectName> is
probably the most similar to the
<objectName  ref="#AlfredJewel">Alfred Jewel</objectName> and was found in
<placeName  ref="#MinsterLovell">Minster
    Lovell</placeName> in
<placeName  ref="#Oxfordshire">Oxfordshire</placeName> and is kept at the
<orgName  ref="#AshmoleanMuseum">Ashmolean Museum</orgName>.
</p>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element objectName
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.personal.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<objectType>** (object type) contains a word or phrase describing the type of object being referred to. [*11.3.2. Material and Object Type*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.canonical
  - *@key*
  - *@ref*

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent

decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The *ref* attribute may be used to point to one or more items within a taxonomy of types of object, defined either internally or externally.

*Example*

```
<physDesc>
  <p> Paper and vellum <objectType>codex</objectType> in modern cloth
binding.</p>
</physDesc>
```

*Example*

```
<physDesc>
  <p>Fragment of a re-used marble <objectType>funerary stele</objectType>.
  </p>
</physDesc>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element objectType
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   macro.phraseSeq}
```

---

## <occupation>

(occupation) contains an informal description of a person's trade, profession or occupation.
[*16.2.2. The Participant Description*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*

    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*

- att.canonical
  - *@key*
  - *@ref*
- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **primary**
**other**
**paid**
**unpaid**

`@scheme` indicates the classification system or taxonomy in use, for example by supplying the identifier of a <taxonomy> element, or pointing to some other resource.
*Status* Optional
*Datatype* teidata.pointer

`@code` identifies an occupation code defined within the classification system or taxonomy defined by the *scheme* attribute.
*Status* Optional
*Datatype* teidata.pointer

*Member of* model.persStateLike
*Contained by*
*namesdates:* person personGrp persona
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound spGrp tech view
*figures:* figure formula notatedMusic table
*gaiji:* g
*header:* biblFull idno
*iso-fs:* fLib fs fvLib
*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The content of this element may be used as an alternative to the more formal specification made possible by its attributes; it may also be used to supplement the formal specification with commentary or clarification.

*Example*

```
<occupation>accountant</occupation>
```

*Example*

```
<occupation  scheme="#occupationtaxonomy"
  code="#acc">accountant</occupation>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element occupation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   attribute scheme { text }?,
   attribute code { text }?,
   macro.specialPara}
```

**<offset>** (offset) marks that part of a relative temporal or spatial expression which indicates the direction of the offset between the two place names, dates, or times involved in the expression. [*14.2.3. Place Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*

    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*

  - att.canonical
    * *@key*
    * *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.offsetLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<placeName key="NRPA1">
  <offset>50 metres below the summit of</offset>
  <geogName>
    <geogFeat>Mount</geogFeat>
    <name>Sinai</name>
  </geogName>
</placeName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element offset
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<opener>** (opener) groups together dateline, byline, salutation, and similar phrases appearing as a preliminary group at the start of a division, especially of a letter. [*4.2. Elements Common to All Divisions*]

*Module* textstructure — Default Text Structure

*Attributes*
- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.written
  - *@hand*

*Member of* model.divTopPart

*Contained by*

*cmc:* post

*core:* lg list

*drama:* castList epilogue performance prologue

*textstructure:* body div div1 div2 div3 div4 div5 div6 div7 group postscript

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice corr date del distinct ellipsis email emph expan foreign gap gb

gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q ref reg rs ruby sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* argument byline dateline epigraph salute signed

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<opener>
  <dateline>Walden, this 29. of August 1592</dateline>
</opener>
```

*Example*

```
<opener>
  <dateline>
    <name  type="place">Great Marlborough Street</name>
    <date>November 11, 1848</date>
  </dateline>
  <salute>My dear Sir,</salute>
</opener>
<p>I am sorry to say that absence from town and other circumstances have
prevented me from
earlier enquiring...</p>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <elementRef key="argument"/>
  <elementRef key="byline"/>
  <elementRef key="dateline"/>
  <elementRef key="epigraph"/>
  <elementRef key="salute"/>
  <elementRef key="signed"/>
```

```
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element opener
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.written.attributes,
   (
      text
    | model.gLike    | model.phrase    | argument    | byline    | dateline    | epigraph
}
```

---

## <oRef>

(orthographic-form reference) in a dictionary example, indicates a reference to the orthographic form(s) of the headword. [*10.4. Headword and Pronunciation References*]

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.lexicographic.normalized
    * *@norm*
    * *@orig*
- att.notated
  - *@notation*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*
- att.typed
  - ~~type~~
  - @subtype

`@type` indicates the kind of typographic modification made to the headword in the reference.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **cap** (capital) indicates first letter is given as capital

**noHyph** (no hyphen) indicates that the headword, though a prefix or suffix, loses its hyphen

*Member of* model.ptrLike.form

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number oRef orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*dictionaries:* oRef

*gaiji:* g

character data

*Example*

```
<entry>
  <form>
    <orth>academy</orth>
  </form>
  <cit  type="example">
    <quote>The Royal <oRef  type="cap"/> of Arts</quote>
  </cit>
</entry>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <elementRef key="oRef"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element oRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.notated.attributes,
   att.pointing.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   ( text | model.gLike | oRef )*
}
```

---

**<org>** (organization) provides information about an identifiable organization such as a business, a tribe, or any other grouping of people. [*14.3.3. Organizational Data*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

  * *@rend*
  * *@style*
  * *@rendition*
  - att.global.responsibility
  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.editLike
  - *@evidence*
  - *@instant*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

`@role` specifies a primary role or classification for the organization.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace

*Note* Values for this attribute may be locally defined by a project, using arbitrary keywords such as artist, employer, familyGroup, or politicalParty, each of which should be associated with a definition. Such local definitions will typically be provided by a <desc> for each <valItem> element in the schema specification of the project's customization.

*Member of* model.personLike

*Contained by*

*corpus:* particDesc

*namesdates:* event listOrg listPerson org

*May contain*

*core:* bibl biblStruct cb desc gb head label lb listBibl milestone name note noteGrp p pb ptr ref rs

*dictionaries:* lang

*header:* biblFull idno

*linking:* ab anchor link linkGrp

*msdescription:* msDesc

*namesdates:* addName bloc climate country district event eventName forename genName geogFeat geogName listEvent listOrg listPerson listPlace location nameLink objectName offset org orgName persName persPronouns person personGrp place placeName population region roleName settlement state surname terrain trait

*tagdocs:* listRef

*transcr:* fw

*Example*

```
<org  xml:id="JAMs">
  <orgName>Justified Ancients of Mummu</orgName>
  <desc>An underground anarchist collective spearheaded by
  <persName>Hagbard Celine</persName>, who fight the Illuminati
     from a golden submarine, the <name>Leif Ericson</name>
```

```
      </desc>
      <bibl>
        <author>Robert Shea</author>
        <author>Robert Anton Wilson</author>
        <title>The Illuminatus! Trilogy</title>
      </bibl>
    </org>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate>
   <classRef key="model.pLike"
    minOccurs="0" maxOccurs="unbounded"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.labelLike"/>
    <classRef key="model.nameLike"/>
    <classRef key="model.placeLike"/>
    <classRef key="model.orgPart"/>
    <classRef key="model.milestoneLike"/>
   </alternate>
  </alternate>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.noteLike"/>
   <classRef key="model.biblLike"/>
   <classRef key="model.ptrLike"/>
   <elementRef key="linkGrp"/>
   <elementRef key="link"/>
  </alternate>
  <classRef key="model.personLike"
   minOccurs="0" maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element org
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.editLike.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   attribute role { list { + } }?,
   (
      model.headLike*,
```

```
      (
         model.pLike*
       | (
            model.labelLike          | model.nameLike          | model.placeLike
      ),
      ( model.noteLike | model.biblLike | model.ptrLike | linkGrp | link )*,
      model.personLike*
   )
}
```

---

## <orgName>

(organization name) contains an organizational name. [*14.2.2. Organizational Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc

- *@generatedBy*

- att.datable
  - *@period*

  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*
    - *@to-custom*
    - *@datingPoint*
    - *@datingMethod*
  - att.datable.iso
    - *@when-iso*
    - *@notBefore-iso*
    - *@notAfter-iso*
    - *@from-iso*
    - *@to-iso*
  - att.datable.w3c
    - *@when*
    - *@notBefore*
    - *@notAfter*
    - *@from*
    - *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.personal
  - *@full*
  - *@sort*

  - att.naming
    - *@role*
    - *@nymRef*

    - att.canonical
      - *@key*
      - *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.nameLike.agent
*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp respStmt rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose setting

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
About a year back, a question of considerable interest was agitated in the
<orgName  key="PAS1"  type="voluntary">
  <placeName  key="PEN">Pennsyla.</placeName> Abolition Society
</orgName> [...]
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element orgName
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.personal.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

## <orig>

<orig> (original form) contains a reading which is marked as following the original, rather than being normalized or corrected. [*3.5.2. Regularization and Normalization 13. Critical Apparatus*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*

*Member of* model.choicePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate

origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example* If all that is desired is to call attention to the original version in the copy text, <orig> may be used alone:

```
<l>But this will be a <orig>meere</orig> confusion</l>
<l>And hardly shall we all be <orig>vnderstoode</orig>
</l>
```

*Example* More usually, an <orig> will be combined with a regularized form within a <choice> element:

```
<l>But this will be a <choice>
   <orig>meere</orig>
   <reg>mere</reg>
  </choice> confusion</l>
<l>And hardly shall we all be <choice>
   <orig>vnderstoode</orig>
   <reg>understood</reg>
  </choice>
</l>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element orig
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   macro.paraContent}
```

---

## <origDate>

<origDate> (origin date) contains any form of date, used to identify the date of origin for a manuscript, manuscript part, or other object. [*11.3.1. Origination*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.calendarSystem
  - *@calendar*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.dimensions

  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
  - att.editLike
    - *@evidence*
    - *@instant*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q ref reg rs ruby sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<origDate notBefore="-0300"
  notAfter="-0200">3rd century BCE</origDate>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element origDate
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
```

```
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.calendarSystem.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   att.typed.attributes,
   ( text | model.gLike | model.phrase | model.global )*
}
```

---

**<origin>** (origin) contains any descriptive or other information concerning the origin of a manuscript, manuscript part, or other object. [*11.8. History*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*

- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.datable
  - *@period*
- att.datable.custom
  - *@when-custom*
  - *@notBefore-custom*
  - *@notAfter-custom*
  - *@from-custom*
  - *@to-custom*
  - *@datingPoint*
  - *@datingMethod*
- att.datable.iso
  - *@when-iso*
  - *@notBefore-iso*
  - *@notAfter-iso*
  - *@from-iso*
  - *@to-iso*
- att.datable.w3c
  - *@when*
  - *@notBefore*
  - *@notAfter*
  - *@from*
  - *@to*

- att.editLike
  - *@evidence*
  - *@instant*

*Contained by:* *msDesc:* history

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<origin notBefore="1802" notAfter="1845"
  evidence="internal" resp="#AMH">Copied in
<name type="origPlace">Derby</name>, probably from an
old Flemish original, between 1802 and 1845, according to
<persName xml:id="AMH">Anne-Mette Hansen</persName>.
</origin>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element origin
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   macro.specialPara}
```

**<origPlace>** (origin place) contains any form of place name, used to identify the place of origin for a manuscript, manuscript part, or other object. [*11.3.1. Origination*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*

    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
    * *@key*
    * *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The *type* attribute may be used to distinguish different kinds of 'origin', for example original place of publication, as opposed to original place of printing.

*Example*

```
<origPlace>Birmingham</origPlace>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
```

```
</content>
```

*Schema Declaration*

```
element origPlace
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<orth>** (orthographic form) gives the orthographic form of a dictionary headword. [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*

* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

• att.lexicographic
– *@expand*
– *@split*
– *@value*
– *@location*
– *@mergedIn*
– *@opt*

– att.datcat
* *@datcat*
* *@valueDatcat*
* *@targetDatcat*

– att.lexicographic.normalized
* *@norm*
* *@orig*

• att.notated
– *@notation*

• att.partials
– *@extent*

• att.typed
– ~~type~~
– @subtype

`@type` gives the type of spelling.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated

*Member of* model.entryPart model.formPart
*Contained by*
*core:* cit
*dictionaries:* dictScrap entryFree form
*namesdates:* nym
*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<form  type="infl">
  <orth>brags</orth>
  <orth>bragging</orth>
  <orth>bragged</orth>
</form>
```

*Example*

```
<form>
  <orth  type="standard"  xml:lang="ko-Hang">치다</orth>
  <orth  type="transliterated"
    xml:lang="ko-Latn">chida</orth>
</form>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element orth
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.notated.attributes,
   att.partials.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.paraContent}
```

---

**<outputRendition>** describes the rendering or appearance intended for all occurrences of an element in a specified context for a specified type of output. [*23.5.4.2. Output Rendition* ]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@scope` provides a way of defining 'pseudo-elements', that is, styling rules applicable to specific sub-portions of an element.
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **first-line** styling applies to the first line of the target element.
**first-letter** styling applies to the first character of the target element.
**before** styling should be applied immediately before the content of the target element.
**after** styling should be applied immediately after the content of the target element.

*Member of* model.oddDecl
*Contained by*
*cmc:* post
*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr
*drama:* camera caption castList epilogue performance prologue set sound tech view
*figures:* cell figDesc figure
*header:* change handNote licence rendition scriptNote tagUsage
*iso-fs:* fDescr fsDescr
*linking:* ab seg
*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote
*namesdates:* occupation
*spoken:* u writing
*tagdocs:* model modelGrp schemaSpec specGrp
*textcrit:* lem rdg witness
*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer
*transcr:* damage metamark mod restore retrace secl supplied surplus
*verse:* rhyme
*May contain* Character data only
*Note* This component of an element specification describes the rendering or appearance intended for all occurrences of the element in a specified context for a specified type of output. The <rendition> element, by contrast, describes the actual rendering or appearance of all occurrences of the specified element in a source document. The *useSourceRendition* attribute may however be used to indicate that formatting information provided by a <rendition> element is to be combined with any provided by an <outputRendition> element.

It is strongly recommended that the W3C Cascading Stylesheet language (CSS2 or later) be used to express the required formatting information.

*Example*

```
<model  predicate="ancestor::p"
  behaviour="inline">
  <outputRendition>font-style: italic;</outputRendition>
</model>
<model  behaviour="block">
  <outputRendition>left-margin: 2em;</outputRendition>
</model>
```

*Content model* `<content> <textNode/></content>`

*Schema Declaration*

```
element outputRendition
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute scope { text }?,
   text
}
```

---

**<p>** (paragraph) marks paragraphs in prose. [*3.1. Paragraphs 7.2.5. Speech Contents*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*

    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.declaring
  - *@decls*
- att.fragmentable
  - *@part*
- att.written
  - *@hand*

*Member of* model.pLike

*Contained by*

*cmc:* post

*core:* item note q quote said sp stage textLang

*corpus:* particDesc setting settingDesc

*drama:* castList epilogue performance prologue set view

*figures:* cell figure

*header:* abstract application availability cRefPattern calendar change correction correspAction correspContext correspDesc editionStmt editorialDecl encodingDesc handNote hyphenation interpretation langUsage licence normalization prefixDef projectDesc publicationStmt punctuation quotation refsDecl samplingDecl scriptNote segmentation seriesStmt sourceDesc stdVals styleDefDecl

*msdescription:* accMat acquisition additional additions binding bindingDesc collation condition custEvent custodialHist decoDesc decoNote filiation foliation handDesc history layout layoutDesc msContents msDesc msFrag msItem msItemStruct msPart musicNotation objectDesc origin physDesc provenance recordHist scriptDesc seal sealDesc signatures source summary support supportDesc surrogates typeDesc typeNote

*namesdates:* climate event langKnowledge listRelation nym object occupation org person personGrp persona place population state terrain trait

*spoken:* broadcast equipment recording recordingStmt scriptStmt transcriptionDesc

*tagdocs:* exemplum remarks specGrp

*textcrit:* lem rdg

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 epigraph front postscript

*transcr:* metamark

*verse:* metDecl

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<p>Hallgerd was outside. <q>There is blood on your axe,</q> she said. <q>What
have you
    done?</q>
</p>
<p>
  <q>I have now arranged that you can be married a second time,</q> replied
Thjostolf.
</p>
<p>
  <q>Then you must mean that Thorvald is dead,</q> she said.
</p>
<p>
  <q>Yes,</q> said Thjostolf. <q>And now you must think up some plan for
me.</q>
</p>
```

*Schematron* <sch:rule context="tei:p"> <sch:report test="(ancestor::tei:ab or ancestor::tei:p) and not( ancestor::tei:floatingText | parent::tei:exemplum | parent::tei:item | parent::tei:note | parent::tei:q | parent::tei:quote | parent::tei:remarks | parent::tei:said | parent::tei:sp | parent::tei:stage | parent::tei:cell | parent::tei:figure )"> Abstract model violation: Paragraphs may not occur inside other paragraphs or ab elements. </sch:report> </sch:rule>

*Schematron* <sch:rule context="tei:l//tei:p"> <sch:assert test="ancestor::tei:floatingText | parent::tei:figure | parent::tei:note"> Abstract model violation: Metrical lines may not contain higher-level structural

elements such as div, p, or ab, unless p is a child of figure or note, or is a descendant of floatingText. </sch:assert> </sch:rule>

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element p
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.declaring.attributes,
   att.fragmentable.attributes,
   att.written.attributes,
   macro.paraContent}
```

---

**<param>** provides a parameter for a model behaviour by supplying its name and an XPath expression identifying the location of its content. [*23.5.4.5. Behaviours and their parameters*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@name` a name for the parameter being supplied.

*Status* Required

*Datatype* teidata.enumerated

*Suggested values include:* **alternate** when used with behaviour `alternate`, a parameter of this name supplies one of the pair of possible values; for example the regularized form rather than the original form within a <choice> element.

**default** when used with behaviour `alternate`, a parameter of this name supplies one of the pair of possible values; for example the original form rather than the regularized form within a <choice> element.

**height** when used with behaviour `graphic`, a parameter of this name supplies a value for the height of the graphic e.g. '300px', '50%'.

**id** a parameter of this name should supply a unique identifier for the element being processed; as for example with the `anchor` behaviour

**label** a parameter of this name should supply an expression to be used to label something, for example `concat('Page ', @n)` for a page beginning or `@n` for a footnote reference; typically used with the `note` or `break` behaviours

**level** when used with the `heading` behaviour, a parameter of this name supplies a positive integer indicating the hierarchic level of a heading.

**link** when used with the `link` behaviour, a parameter of this name should supply a URL to be used as the target of a link.

**place** when used with the `note` behaviour, a parameter of this name should provide a string which describes the intended placement of some text; typical values include 'margin', 'footnote', 'endnote', 'inline', 'bottom'

**type** a parameter of this name can be used to categorize the specified behaviour in any way; for example the kind of break (when used with the `break` behaviour) or the kind of index to be generated (if used with the `index` behaviour) etc.

**url** when used with behaviour `graphic`, a parameter of this name supplies a a URL indicating the graphic intended.

**width** when used with behaviour `graphic`, a parameter of this name supplies a value for the width of the graphic e.g. '400px', '70%'.

`@value` supplies an XPath expression which when evaluated provides the value for the parameter

*Status* Required

*Datatype* teidata.xpath

*Contained by:* tei.gloss model

*May contain* Empty element

*Note* An implementation may require parameters other than those listed here, and is not required to follow this partial specification.

The names and datatypes of the expected parameters should be documented in the corresponding customization using a <paramSpec> element. Literal strings provided in an XPath expression should be quoted.

*Example* In this example, which will be processed for a <choice> element which has both <sic> and <corr> child elements, the `default` parameter will provide the value of the child <corr> element, and the `alternate` parameter will provide that of the child <sic> elements. If neither <param> element was supplied, both elements would still be available to an application, but the application would need to distinguish them for itself.

```
<elementSpec ident="choice">
  <model predicate="sic and corr"
    behaviour="alternate">
    <param name="default" value="corr"/>
    <param name="alternate" value="sic"/>
  </model>
</elementSpec>
```

*Example*

```
<elementSpec ident="graphic" mode="change">
  <model behaviour="graphic">
    <param name="url" value="@url"/>
    <param name="width" value="@width"/>
    <param name="height" value="@height"/>
  </model>
</elementSpec>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element param
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute name
   {
      "alternate"
    | "default"
    | "height"
    | "id"
    | "label"
    | "level"
    | "link"
    | "place"
    | "type"
    | "url"
    | "width"
   },
   attribute value { text },
```

```
   empty
}
```

---

**<paramList>** list of parameter specifications. [*23.5.4.5. Behaviours and their parameters 23.5.4.8. Defining a processing model*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by* *tagdocs:* valItem

*May contain*

*tagdocs:* paramSpec

*Note* The <paramList> element provides a a mechanism to document parameter specifications using child <paramSpec> elements.

*Example*

```
<valItem  ident="alternate">
  <desc  versionDate="2015-08-21"
    xml:lang="en">create a specialized display of alternating elements for
displaying the preferred version and an alternative, both at once or by some
method of toggling between the two.</desc>
  <paramList>
    <paramSpec  ident="default">
      <desc>preferred content</desc>
    </paramSpec>
    <paramSpec  ident="alternate">
      <desc>alternate content</desc>
    </paramSpec>
  </paramList>
</valItem>
```

*Content model*

```
<content>
 <elementRef minOccurs="0"
  maxOccurs="unbounded" key="paramSpec"/>
</content>
```

*Schema Declaration*

```
element paramList
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   paramSpec*
}
```

---

## <paramSpec> supplies specification for one parameter of a model behaviour. [*23.5.4.8. Defining a processing model*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
- att.global.analytic
  - *@ana*

- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.identified
  - *@ident*
  - *@predeclare*
  - *@module*

- att.combinable
  - *@mode*

- att.deprecated
  - *@validUntil*

*Contained by* *tagdocs:* paramList

*May contain*

*core:* desc gloss

*tagdocs:* equiv

*Note* Where a model behaviour uses more than one parameter, individual <paramSpec> elements should be grouped together using a <paramList> element, as above. Parameter specifications are provided within the <valItem> used to define a particular behaviour, which forms part of the specification of a <model> element's *behaviour* attribute. In the example above, the behaviour `link` has two parameters: content and link.

Using a <desc> element within a <paramSpec> is optional but recommended practice.

*Example*

```
<valItem  ident="link">
  <desc>create a hyperlink</desc>
  <paramList>
    <paramSpec  ident="content">
      <desc>supplies the location of some content describing the link</desc>
    </paramSpec>
    <paramSpec  ident="link">
      <desc>supplies the location of the intended URL</desc>
    </paramSpec>
  </paramList>
</valItem>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.identEquiv"/>
  <classRef key="model.descLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element paramSpec
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.identified.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   ( model.identEquiv | model.descLike )*
}
```

---

**<particDesc>** (participation description) describes the identifiable speakers, voices, or other participants in any kind of text or other persons named or otherwise referred to in a text, edition, or metadata. [*16.2. Contextual Information*]

*Module* corpus — Language Corpora

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.declarable
  - *@default*

*Member of* model.profileDescPart

*Contained by*

*header:* profileDesc

*May contain*

*core:* p

*linking:* ab

*namesdates:* listOrg listPerson org person personGrp

*Note* May contain a prose description organized as paragraphs, or a structured list of persons and person groups, with an optional formal specification of any relationships amongst them.

*Example*

```
<particDesc>
 <listPerson>
   <person xml:id="P-1234" sex="2" age="mid">
     <p>Female informant, well-educated, born in
          Shropshire UK, 12 Jan 1950, of unknown occupation. Speaks French
fluently.
          Socio-Economic status B2.</p>
   </person>
   <person xml:id="P-4332" sex="1">
```

```
      <persName>
        <surname>Hancock</surname>
        <forename>Antony</forename>
        <forename>Aloysius</forename>
        <forename>St John</forename>
      </persName>
      <residence  notAfter="1959">
        <address>
          <street>Railway Cuttings</street>
          <settlement>East Cheam</settlement>
        </address>
      </residence>
      <occupation>comedian</occupation>
    </person>
    <listRelation>
      <relation  type="personal"  name="spouse"
        mutual="#P-1234 #P-4332"/>
    </listRelation>
  </listPerson>
</particDesc>
```

This example shows both a very simple person description, and a very detailed one, using some of the more specialized elements from the module for Names and Dates.

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde"  value="tei:particDesc"/> </sch:pattern>

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.personLike"/>
   <elementRef key="listPerson"/>
   <elementRef key="listOrg"/>
  </alternate>
 </alternate>
</content>
```

*Schema Declaration*

```
element particDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   ( model.pLike+ | ( model.personLike | listPerson | listOrg )+ )
}
```

**<path>** (path) defines any line passing through two or more points within a <surface> element. [*12.1. Digital Facsimiles 12.2.2. Embedded Transcription*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*
- att.coordinated
  - ~~points~~
  - @start

- @ulx
- @uly
- @lrx
- @lry

`@points` identifies a line within the container or bounding box specified by the parent element by means of a series of two or more pairs of numbers, each of which gives the x,y coordinates of a point on the line.
*Derived from* att.coordinated
*Status* Optional
*Datatype* 2–∞ occurrences of teidata.point separated by whitespace

*Member of* model.linePart
*Contained by*
*transcr:* line surface zone
*May contain* Empty element
*Note* Although the simplest form of a path is a straight line between two points, a line with more than two points may bend at any point. The order of coordinates in *points* is significant, because the line follows the coordinate sequence.

To specify a closed polygon, use the <zone> element rather than the <path> element.
*Example*

```
<surface ulx="0" uly="0" lrx="443" lry="272">
  <graphic url="facs-fig3.jpg"/>
  <path points="74,73 171,244"/>
  <path points="71,203 173,116"/>
</surface>
```

*Schematron* Since a <path> represents a line with distinct start and end points, the last coordinate should not be the same as the first coordinate. <sch:rule context="tei:path[@points]"> <sch:let name="firstPair" value="tokenize( normalize-space( @points ), ' ')[1]"/> <sch:let name="lastPair" value="tokenize( normalize-space( @points ), ' ')[last()]"/> <sch:let name="firstX" value="xs:float( substring-before( $firstPair, ',') )"/> <sch:let name="firstY" value="xs:float( substring-after( $firstPair, ',') )"/> <sch:let name="lastX" value="xs:float( substring-before( $lastPair, ',') )"/> <sch:let name="lastY" value="xs:float( substring-after( $lastPair, ',') )"/> <sch:report test="$firstX eq $lastX and $firstY eq $lastY">The first and last elements of this path are the same. To specify a closed polygon, use the zone element rather than the path element. </sch:report> </sch:rule>

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element path
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   att.written.attributes,
   att.coordinated.attribute.start,
   att.coordinated.attribute.ulx,
   att.coordinated.attribute.uly,
```

```
   att.coordinated.attribute.lrx,
   att.coordinated.attribute.lry,
   attribute points { list { * } }?,
   empty
}
```

---

**<pause>** (pause) marks a pause either between or within utterances. [*8.3.2. Pausing*]

*Module* spoken — Transcriptions of Speech

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.ascribed.directed
  - *@toWhom*
  - att.ascribed

    * *@who*
  - att.cmc
    – *@generatedBy*
  - att.timed
    – *@start*
    – *@end*

    – att.duration
    * att.duration.iso
    · *@dur-iso*
    * att.duration.w3c
    · *@dur*
  - att.typed
    – *@type*
    – *@subtype*

*Member of* model.global.spoken

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* annotationBlock u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Example*

```
<pause dur="PT42S" type="pregnant"/>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element pause
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.directed.attributes,
   att.ascribed.attributes,
   att.cmc.attributes,
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   att.typed.attributes,
   empty
}
```

---

**<pb>** (page beginning) marks the beginning of a new page in a paginated document. [*3.11.3. Milestone Elements*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*

    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.breaking
  - *@break*
- att.cmc
  - *@generatedBy*
- att.edition
  - *@ed*
  - *@edRef*
- att.spanning
  - *@spanTo*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.milestoneLike

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list listBibl measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent

decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc subst supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Note* A <pb> element should appear at the start of the page which it identifies. The global *n* attribute indicates the number or other value associated with this page. This will normally be the page number or signature printed on it, since the physical sequence number is implicit in the presence of the <pb> element itself.

The *type* attribute may be used to characterize the page beginning in any respect. The more specialized attributes *break*, *ed*, or *edRef* should be preferred when the intent is to indicate whether or not the page beginning is word-breaking, or to note the source from which it derives.

*Example* Page numbers may vary in different editions of a text.

```
<p> ... <pb  n="145"  ed="ed2"/>
<!-- Page 145 in edition "ed2" starts here --> ... <pb  n="283"  ed="ed1"/>
<!-- Page 283 in edition "ed1" starts here--> ... </p>
```

*Example* A page beginning may be associated with a facsimile image of the page it introduces by means of the *facs* attribute

```
<body>
  <pb  n="1"  facs="page1.png"/>
<!-- page1.png contains an image of the page;
the text it contains is encoded here -->
  <p>
<!-- ... -->
  </p>
  <pb  n="2"  facs="page2.png"/>
<!-- similarly, for page 2 -->
  <p>
<!-- ... -->
  </p>
</body>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element pb
{
   att.global.attributes,
   att.global.analytic.attributes,
```

```
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.breaking.attributes,
   att.cmc.attributes,
   att.edition.attributes,
   att.spanning.attributes,
   att.typed.attributes,
   empty
}
```

---

**<pc>** (punctuation character) contains a character or string of characters regarded as constituting a single punctuation mark. [*18.1.2. Below the Word Level 18.4.2. Lightweight Linguistic Annotation*]

*Module* analysis — Simple Analytic Mechanisms

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*

  * *@resp*
  - att.global.source
  * *@source*
- att.cmc
  - *@generatedBy*
- att.linguistic
  - *@lemma*
  - *@lemmaRef*
  - *@pos*
  - *@msd*
  - *@join*
  - att.lexicographic.normalized
  * *@norm*
  * *@orig*
- att.segLike
  - *@function*
  - att.datcat
  * *@datcat*
  * *@valueDatcat*
  * *@targetDatcat*
  - att.fragmentable
  * *@part*
  - att.metrical
  * *@met*
  * *@real*
  * *@rhyme*
- att.typed
  - *@type*
  - *@subtype*

`@force` indicates the extent to which this punctuation mark conventionally separates words or phrases.
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **strong** the punctuation mark is a word separator
**weak** the punctuation mark is not a word separator
**inter** the punctuation mark may or may not be a word separator

`@unit` provides a name for the kind of unit delimited by this punctuation mark.
*Status* Optional
*Datatype* teidata.enumerated

`@pre` indicates whether this punctuation mark precedes or follows the unit it delimits.
*Status* Optional
*Datatype* teidata.truthValue

*Member of* model.linePart model.segLike

*Contained by*

*analysis:* cl phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain*

*analysis:* c

*core:* abbr add choice corr del expan orig reg sic unclear

*gaiji:* g

*transcr:* am damage ex handShift mod redo restore retrace secl subst supplied surplus undo

character data

*Example*

```
<phr>
  <w>do</w>
  <w>you</w>
  <w>understand</w>
  <pc  type="interrogative">?</pc>
</phr>
```

*Example* Example encoding of the German sentence *Wir fahren in den Urlaub.*, encoded with attributes from `att.linguistic` discussed in section *18.4.2. Lightweight Linguistic Annotation*.

```
<s>
  <w  pos="PPER"  msd="1.Pl.*.Nom">Wir</w>
  <w  pos="VVFIN"  msd="1.Pl.Pres.Ind">fahren</w>
  <w  pos="APPR"  msd="--">in</w>
  <w  pos="ART"  msd="Def.Masc.Akk.Sg.">den</w>
```

```
      <w  pos="NN"  msd="Masc.Akk.Sg.">Urlaub</w>
      <pc  pos="$."  msd="--"  join="left">.</pc>
    </s>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <elementRef key="c"/>
  <classRef key="model.pPart.edit"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element pc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.linguistic.attributes,
   att.lexicographic.normalized.attributes,
   att.segLike.attributes,
   att.datcat.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   attribute force { "strong" | "weak" | "inter" }?,
   attribute unit { text }?,
   attribute pre { text }?,
   ( text | model.gLike | c | model.pPart.edit )*
}
```

---

**<per>** (person) contains an indication of the grammatical person (1st, 2nd, 3rd, etc.) associated with a given inflected form in a dictionary. [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.lexicographic.normalized
    * *@norm*
    * *@orig*

*Member of* model.morphLike
*Contained by*
*dictionaries:* dictScrap entryFree form gramGrp
*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* This element is synonymous with <gram type="person">.

*Example* Taken from *Wörterbuch der Deutschen Sprache. Veranstaltet und herausgegeben von Joachim Heinrich Campe. Vierter Theil. S - bis - T. (Braunschweig 1810. In der Schulbuchhandlung)*:

> Treffen, v. unregelm. … du triffst, …

```
<entry>
  <form type="inflected">
    <gramGrp>
      <per value="2"/>
      <number value="singular"/>
      <tns value="present"/>
      <mood value="indicative"/>
    </gramGrp>
    <form type="personalpronoun">
      <orth>du</orth>
    </form>
    <form type="headword">
      <orth>
        <oRef>triffst</oRef>
      </orth>
    </form>
```

```
      </form>
    </entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element per
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   macro.paraContent}
```

---

## <performance>

<performance> (performance) contains a section of front or back matter describing how a dramatic piece is to be performed in general or how it was performed on some specific occasion. [*7.1.3. Records of Performances 7.1. Front and Back Matter* ]

*Module* drama — Performance Texts

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Member of* model.frontPart.drama

*Contained by*

*textcrit:* lem rdg

*textstructure:* back front

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct cb cit desc ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* contains paragraphs and an optional cast list only.

*Example*

```
<performance>
  <p>
    <rs  type="place">Gateway Theatre, Edinburgh</rs>, <date>6 September
```

```
1948</date>
   <castList>
     <castItem>
       <role>Anath Bithiah</role>
       <actor>Athene Seyler</actor>
     </castItem>
     <castItem>
       <role>Shendi</role>
       <actor>Robert Rietty</actor>
     </castItem>
   </castList>
  </p>
  <p>Directed by <name>E. Martin Browne</name>
  </p>
</performance>
```

*Example*

```
<performance>
  <p>Cast of the original production at the <rs  type="place">Savoy Theatre,
London,</rs> on
  <date>September 24, 1907</date>
    <castList>
      <castItem>Colonel Hope : Mr A.E.George</castItem>
    </castList>
  </p>
</performance>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.common"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divBottom"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element performance
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      ( model.divTop | model.global )*,
      ( ( model.common, model.global* )+ ),
      ( ( model.divBottom, model.global* )* )
   )
}
```

---

**<persName>** (personal name) contains a proper noun or proper-noun phrase referring to a person, possibly including one or more of the person's forenames, surnames, honorifics, added names, etc. [*14.2.1. Personal Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*
  - att.datable.custom
  * *@when-custom*
  * *@notBefore-custom*
  * *@notAfter-custom*
  * *@from-custom*
  * *@to-custom*
  * *@datingPoint*
  * *@datingMethod*
  - att.datable.iso
  * *@when-iso*
  * *@notBefore-iso*
  * *@notAfter-iso*
  * *@from-iso*
  * *@to-iso*
  - att.datable.w3c
  * *@when*
  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.personal
  - *@full*
  - *@sort*
  - att.naming
  * *@role*
  * *@nymRef*
  * att.canonical
  · *@key*
  · *@ref*

- att.typed
  - *@type*
  - *@subtype*

*Member of* model.nameLike.agent model.persStateLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp respStmt rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose setting

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<persName>
  <forename>Edward</forename>
  <forename>George</forename>
  <surname  type="linked">Bulwer-Lytton</surname>, <roleName>Baron Lytton of
  <placeName>Knebworth</placeName>
  </roleName>
</persName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element persName
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.personal.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

**<person>** (person) provides information about an identifiable individual, for example a participant in a language interaction, or a person referred to in a historical source. [*14.3.2. The Person Element 16.2.2. The Participant Description*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.editLike
  - *@evidence*
  - *@instant*
- att.sortable
  - *@sortKey*

`@role` specifies a primary role or classification for the person.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace

*Note* Values for this attribute may be locally defined by a project, using arbitrary keywords such as artist, employer, author, relative, or servant, each of which should

be associated with a definition. Such local definitions will typically be provided by a <valList> element in the project schema specification.

`@sex` specifies the sex of the person.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.sex separated by whitespace

*Note* Values for this attribute may be defined locally by a project, or they may refer to an external standard.

`@gender` specifies the gender of the person.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.gender separated by whitespace

*Note* Values for this attribute may be defined locally by a project, or they may refer to an external standard.

`@age` specifies an age group for the person.

*Status* Optional

*Datatype* teidata.enumerated

*Note* Values for this attribute may be locally defined by a project, using arbitrary keywords such as infant, child, teen, adult, or senior, each of which should be associated with a definition. Such local definitions will typically be provided by a <valList> element in the project schema specification.

*Member of* model.personLike

*Contained by*

*corpus:* particDesc

*namesdates:* event listPerson org

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* bibl biblStruct cb ellipsis gap gb index lb listBibl milestone name note noteGrp p pb ptr ref

*figures:* figure notatedMusic

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* affiliation age birth death education event faith floruit gender langKnowledge listEvent nationality occupation persName persPronouns persona residence sex socecStatus state trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* listRef

*textcrit:* app witDetail

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* May contain either a prose description organized as paragraphs, or a sequence of more specific demographic elements drawn from the `model.personPart` class.

*Example*

```
<person sex="F" age="adult">
  <p>Female respondent, well-educated, born in Shropshire UK, 12 Jan 1950, of
unknown occupation. Speaks French fluently. Socio-Economic
    status B2.</p>
</person>
```

*Example*

```
<person sex="intersex" role="god"
  age="immortal">
  <persName>Hermaphroditos</persName>
  <persName xml:lang="grc">Ἑρμαφρόδιτος</persName>
</person>
```

*Example*

```
<person xml:id="Ovi01" sex="M" role="poet">
  <persName xml:lang="en">Ovid</persName>
  <persName xml:lang="la">Publius Ovidius Naso</persName>
  <birth when="-0044-03-20"> 20 March 43 BC <placeName>
     <settlement type="city">Sulmona</settlement>
     <country key="IT">Italy</country>
    </placeName>
  </birth>
  <death notBefore="0017" notAfter="0018">17 or 18 AD <placeName>
     <settlement type="city">Tomis (Constanta)</settlement>
     <country key="RO">Romania</country>
    </placeName>
  </death>
</person>
```

*Example* The following exemplifies an adaptation of the vCard standard to indicate an unknown gender for a fictional character.

```
<person xml:id="ariel" gender="U">
  <persName>Ariel</persName>
  <note>Character in <title level="m">The Tempest</title>.</note>
</person>
```

*Example* This example demonstrates the use of a <ref> element to provide more information about a person.

```
<person age="G2" role="author"
  xml:id="W0212" sex="F">
  <birth when="1787"/>
  <death when="1855"/>
  <persName type="main">Mitford, Mary Russell (1787–1855)</persName>
  <persName resp="#Nicoll">MITFORD, MARY RUSSELL</persName>
  <listBibl type="lacyTitles">
    <desc>Lacy's Acting Editions</desc>
    <bibl>
      <ref target="lacy:L1280">Foscari</ref>
    </bibl>
    <bibl>
      <ref target="lacy:L1337">Rienzi</ref>
    </bibl>
  </listBibl>
  <listRef type="seeAlso">
    <ref target="https://www.victorianresearch.org/atcl/show_author.php?aid=1386">ATCL</ref>
    <ref target="https://doi.org/10.1093/ref:odnb/18859">ODNB</ref>
    <ref target="https://en.wikipedia.org/wiki/Mary_Russell_Mitford">Wikipedia</ref>
    <ref target="https://digitalmitford.org">Digital Mitford</ref>
  </listRef>
</person>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.personPart"/>
   <classRef key="model.global"/>
   <classRef key="model.ptrLike"/>
  </alternate>
 </alternate>
</content>
```

*Schema Declaration*

```
element person
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.editLike.attributes,
   att.sortable.attributes,
   attribute role { list { + } }?,
   attribute sex { list { + } }?,
   attribute gender { list { + } }?,
   attribute age { text }?,
   ( model.pLike+ | ( model.personPart | model.global | model.ptrLike )* )
}
```

---

**<persona>** provides information about one of the personalities identified for a given individual, where an individual has multiple personalities. [*14.3.2. The Person Element*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*

  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.editLike
  - *@evidence*
  - *@instant*
- att.sortable
  - *@sortKey*

`@role` specifies a primary role or classification for the persona.
: *Status* Optional
: *Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace
: *Note* Values for this attribute may be locally defined by a project, using arbitrary keywords such as artist, employer, author, relative, or servant, each of which should be associated with a definition. Such local definitions will typically be provided by a <valList> element in the project schema specification.

`@sex` specifies the sex of the persona.
: *Status* Optional
: *Datatype* 1–∞ occurrences of teidata.sex separated by whitespace
: *Note* Values for this attribute may be locally defined by a project, or may refer to an external standard.

`@gender` specifies the gender of the persona.
: *Status* Optional
: *Datatype* 1–∞ occurrences of teidata.gender separated by whitespace
: *Note* Values for this attribute may be locally defined by a project, or they may refer to an external standard.

`@age` specifies an age group for the persona.
: *Status* Optional
: *Datatype* teidata.enumerated
: *Note* Values for this attribute may be locally defined by a project, using arbitrary keywords such as infant, child, teen, adult, or senior, each of which should be

associated with a definition. Such local definitions will typically be provided by a <valList> element in the project schema specification.

*Member of* model.persStateLike

*Contained by*

*namesdates:* person personGrp persona

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* bibl biblStruct cb ellipsis gap gb index lb listBibl milestone name note noteGrp p pb

*figures:* figure notatedMusic

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* affiliation age birth death education event faith floruit gender langKnowledge listEvent nationality occupation persName persPronouns persona residence sex socecStatus state trait

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* Note that a persona is not the same as a role. A role may be assumed by different people on different occasions, whereas a persona is unique to a particular person, even though it may resemble others. Similarly, when an actor takes on or enacts the role of a historical person, they do not thereby acquire a new persona.

*Example*

```
<person sex="M" age="adult">
  <persona sex="M">
    <persName>Dr Henry Jekyll</persName>
  </persona>
  <persona sex="M" age="youth">
    <persName>Edward Hyde</persName>
  </persona>
</person>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.personPart"/>
   <classRef key="model.global"/>
  </alternate>
 </alternate>
</content>
```

*Schema Declaration*

```
element persona
{
```

```
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.editLike.attributes,
   att.sortable.attributes,
   attribute role { list { + } }?,
   attribute sex { list { + } }?,
   attribute gender { list { + } }?,
   attribute age { text }?,
   ( model.pLike+ | ( model.personPart | model.global )* )
}
```

---

**<personGrp>** (personal group) describes a group of individuals treated as a single person for analytic purposes. [*16.2.2. The Participant Description*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*

- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.sortable
  - *@sortKey*

`@role` specifies the role of this group of participants in the interaction.
*Status* Optional
*Datatype* teidata.enumerated
*Note* Values for this attribute may be locally defined by a project, using arbitrary keywords such as movement, employers, relatives, or servants, each of which should be associated with a definition. Such local definitions will typically be provided by a <valList> element in the project schema specification.

`@sex` specifies the sex of the participant group.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.sex separated by whitespace
*Note* Values for this attribute may be defined locally by a project, or they may refer to an external standard.

`@gender` specifies the gender of the participant group.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.gender separated by whitespace
*Note* Values for this attribute may be defined locally by a project, or they may refer to an external standard.

`@age` specifies the age group of the participants.
*Status* Optional
*Datatype* teidata.enumerated
*Note* Values for this attribute may be locally defined by a project, using arbitrary keywords such as infant, child, teen, adult, or senior, each of which should be associated with a definition. Such local definitions will typically be provided by a <valList> element in the project schema specification.

`@size` describes informally the size or approximate size of the group for example by means of a number and an indication of accuracy e.g. approx 200.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

*Member of* model.personLike
*Contained by*
*corpus:* particDesc
*namesdates:* event listPerson org
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* bibl biblStruct cb ellipsis gap gb index lb listBibl milestone name note noteGrp p pb
*figures:* figure notatedMusic
*header:* biblFull idno
*iso-fs:* fLib fs fvLib
*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* msDesc

*namesdates:* affiliation age birth death education event faith floruit gender langKnowledge listEvent nationality occupation persName persPronouns persona residence sex socecStatus state trait

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* May contain a prose description organized as paragraphs, or any sequence of demographic elements in any combination.

The global *xml:id* attribute should be used to identify each speaking participant in a spoken text if the *who* attribute is specified on individual utterances.

*Example*

```
<personGrp  xml:id="pg1"  role="audience"
  sex="mixed"  size="approx 50"/>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.personPart"/>
   <classRef key="model.global"/>
  </alternate>
 </alternate>
</content>
```

*Schema Declaration*

```
element personGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.sortable.attributes,
   attribute role { text }?,
   attribute sex { list { + } }?,
   attribute gender { list { + } }?,
   attribute age { text }?,
   attribute size { list { + } }?,
   ( model.pLike+ | ( model.personPart | model.global )* )
}
```

---

## <persPronouns>

(personal pronouns) indicates the personal pronouns used, or assumed to be used, by the individual being described. [*14.3.2.1. Personal Characteristics*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes* • att.global

  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*

- att.datable.iso
  - *@when-iso*
  - *@notBefore-iso*
  - *@notAfter-iso*
  - *@from-iso*
  - *@to-iso*
- att.datable.w3c
  - *@when*
  - *@notBefore*
  - *@notAfter*
  - *@from*
  - *@to*
- att.typed
  - *@type*
  - *@subtype*

`@evidence` (evidence) indicates support for the listed personal pronouns.

*Status* Recommended

*Datatype* teidata.enumerated

*Suggested values include:* **conjecture** (conjecture) The given value was selected based on assumptions by someone besides the person to whom this pronoun applies. As a result, the value may be erroneous.

**selfIdentification** (self identification) The given value has been explicitly stated or confirmed by the person to whom this pronoun applies.

**trustedThirdParty** (trusted third party) The given value has been supplied by another individual trusted by the encoder to know the preferences of the person to whom this pronoun applies.

`@value` (value) supplies a regularized value for personal pronouns.

*Status* Recommended

*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace

*Sample values include:* **e** (e) e/eirs

**he** (he) he/him/his

**she** (she) she/her/hers

**they** (they) they/them/theirs

*Note* The sample values list shown here is intended to be reflective of English usage. There is nothing restricting users from defining lists reflective of other languages, e.g. elle, il, and ils.

*Member of* model.persNamePart model.persStateLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<person>
  <persName>
    <forename>SUE</forename>
    <addName>the T. rex</addName>
  </persName>
  <residence>The Field Museum. Chicago, Illinois, United States.</residence>
  <sex  value="0"/>
  <persPronouns  value="they">they/them</persPronouns>
  <note>
    <cit>
      <quote>Specimen FMNH PR 2081. Legendary Fossil. Apex
            Predator. National Treasure. <emph  style="text-transform:uppercase;
letter-spacing:0.25rem;">Murderbird.</emph>
      </quote>
      <bibl>SUEtheTrex, Twitter biography.
      <ptr  target="https://twitter.com/SUEtheTrex"/>.
            Accessed <date  when="2020-03-25">March 25th, 2020</date>.</bibl>
    </cit>
  </note>
</person>
```

*Example*

```
<docAuthor>
  <persName>Lal Zimman</persName>
  <persPronouns  value="he">(he/him/his)</persPronouns>
  <ref  target="#Name">(FAQ)</ref>
  <persName  type="IPA">[lɑɬ 'zimn̩]</persName>
  <email>zimman at ucsb dot edu</email>
  <roleName>Assistant Professor of Linguistics</roleName>
  <roleName>Affiliated Faculty in Feminist Studies</roleName>
  <address>
    <addrLine>South Hall 3518</addrLine>
    <addrLine>University of California, Santa Barbara</addrLine>
  </address>
</docAuthor>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element persPronouns
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
```

```
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.typed.attributes,
   attribute evidence
   {
      "conjecture" | "selfIdentification" | "trustedThirdParty"
   }?,
   attribute value { list { + } }?,
   macro.phraseSeq}
```

---

**<phr>** (phrase) represents a grammatical phrase. [*18.1. Linguistic Segment Categories*]

*Module* analysis — Simple Analytic Mechanisms

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source

    * *@source*
  - att.cmc
    – *@generatedBy*
  - att.notated
    – *@notation*
  - att.segLike
    – *@function*

    – att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
    – att.fragmentable
    * *@part*
    – att.metrical
    * *@met*
    * *@real*
    * *@rhyme*
  - att.typed
    – *@type*
    – *@subtype*

*Member of* model.segLike

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The *type* attribute may be used to indicate the type of phrase, taking values such as noun, verb, preposition, etc. as appropriate.

*Example*

```
<phr  type="verb"
  function="extraposted_modifier">To talk
<phr  type="preposition"
    function="complement">of
  <phr  type="noun"  function="object">many things</phr>
  </phr>
</phr>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element phr
{
   att.global.attributes,
   att.global.analytic.attributes,
```

```
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.notated.attributes,
   att.segLike.attributes,
   att.datcat.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<physDesc>** (physical description) contains a full physical description of a manuscript, manuscript part, or other object optionally subdivided using more specialized elements from the `model.physDescPart` class. [*11.7. Physical Description*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

*Contained by:* *msDescription:* msDesc msFrag msPart
*namesdates:* object
*May contain*
*core:* p
*linking:* ab
*msdescription:* accMat additions bindingDesc decoDesc handDesc musicNotation objectDesc scriptDesc sealDesc typeDesc
*Example*

```
<physDesc>
  <objectDesc  form="codex">
    <supportDesc  material="perg">
      <support>Parchment.</support>
      <extent>i + 55 leaves
      <dimensions  scope="all"  type="leaf"
          unit="inch">
          <height>7¼</height>
          <width>5⅜</width>
        </dimensions>
      </extent>
    </supportDesc>
    <layoutDesc>
      <layout  columns="2">In double columns.</layout>
    </layoutDesc>
  </objectDesc>
  <handDesc>
    <p>Written in more than one hand.</p>
  </handDesc>
  <decoDesc>
    <p>With a few coloured capitals.</p>
  </decoDesc>
</physDesc>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.pLike" minOccurs="0"
   maxOccurs="unbounded"/>
  <classRef key="model.physDescPart"
   expand="sequenceOptional"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element physDesc
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      model.pLike*,
      objectDesc?,
      handDesc?,
      typeDesc?,
      scriptDesc?,
      musicNotation?,
      decoDesc?,
      additions?,
      bindingDesc?,
      sealDesc?,
      accMat?
   )
}
```

---

**<place>** (place) contains data about a geographic location. [*14.3.4. Places*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.editLike
  - *@evidence*
  - *@instant*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.placeLike

*Contained by*

*corpus:* settingDesc

*namesdates:* event listPlace org place

*May contain*

*core:* bibl biblStruct desc head label listBibl name note noteGrp p ptr ref

*header:* biblFull idno

*linking:* ab link linkGrp

*msdescription:* msDesc

*namesdates:* bloc climate country district event geogName listEvent listPlace location place placeName population region settlement state terrain trait

*tagdocs:* listRef

*Example*

```
<place>
  <country>Lithuania</country>
  <country  xml:lang="lt">Lietuva</country>
  <place>
    <settlement>Vilnius</settlement>
  </place>
  <place>
    <settlement>Kaunas</settlement>
  </place>
</place>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate>
   <classRef key="model.pLike"
    minOccurs="0" maxOccurs="unbounded"/>
```

```
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.labelLike"/>
    <classRef key="model.placeStateLike"/>
    <classRef key="model.eventLike"/>
    <elementRef key="name"/>
   </alternate>
  </alternate>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.noteLike"/>
   <classRef key="model.biblLike"/>
   <classRef key="model.ptrLike"/>
   <elementRef key="idno"/>
   <elementRef key="linkGrp"/>
   <elementRef key="link"/>
  </alternate>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.placeLike"/>
   <elementRef key="listPlace"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element place
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.editLike.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   (
      model.headLike*,
      (
         model.pLike*
       | ( model.labelLike | model.placeStateLike | model.eventLike | name )*
      ),
      (
         model.noteLike       | model.biblLike       | model.ptrLike       | idno       |
      ( model.placeLike | listPlace )*
   )
}
```

---

## <placeName>

(place name) contains an absolute or relative place name. [*14.2.3. Place Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*
    - *@to-custom*
    - *@datingPoint*

    * *@datingMethod*
    – att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
    – att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
  - att.editLike
    – *@evidence*
    – *@instant*
  - att.personal
    – *@full*
    – *@sort*
    – att.naming
    * *@role*
    * *@nymRef*
    * att.canonical
    · *@key*
    · *@ref*
  - att.typed
    – *@type*
    – *@subtype*

*Member of* model.placeNamePart model.settingPart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose setting

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage unitDef

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions altIdentifier catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msIdentifier musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectIdentifier objectName occupation offset org orgName persName persPronouns place placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<placeName>
  <settlement>Rochester</settlement>
```

```
  <region>New York</region>
</placeName>
```

*Example*

```
<placeName>
  <geogName>Arrochar Alps</geogName>
  <region>Argylshire</region>
</placeName>
```

*Example*

```
<placeName>
  <measure>10 miles</measure>
  <offset>Northeast of</offset>
  <settlement>Attica</settlement>
</placeName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element placeName
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.personal.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

## <population>

(population) contains information about the population of a place. [*14.3.4.3. States, Traits, and Events*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*

  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso

    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*
  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
    * *@key*
    * *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.placeStateLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num

orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns place placeName population region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*certainty:* precision

*core:* bibl biblStruct desc head label listBibl note noteGrp p

*header:* biblFull

*linking:* ab

*msdescription:* msDesc

*namesdates:* population

*Example*

```
<population when="2001-04" resp="#UKCensus">
  <population type="white">
    <desc>54153898</desc>
  </population>
  <population type="asian">
    <desc>11811423</desc>
  </population>
  <population type="black">
    <desc>1148738</desc>
  </population>
  <population type="mixed">
    <desc>677117</desc>
  </population>
  <population type="chinese">
    <desc>247403</desc>
  </population>
```

```
      <population  type="other">
        <desc>230615</desc>
      </population>
    </population>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="precision" minOccurs="0"
   maxOccurs="unbounded"/>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <sequence minOccurs="0">
   <alternate>
    <classRef key="model.pLike"
     minOccurs="1" maxOccurs="unbounded"/>
    <classRef key="model.labelLike"
     minOccurs="1" maxOccurs="unbounded"/>
   </alternate>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.noteLike"/>
    <classRef key="model.biblLike"/>
   </alternate>
  </sequence>
  <elementRef key="population"
   minOccurs="0" maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element population
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   (
      precision*,
      model.headLike*,
```

```
      (
         (
            ( model.pLike+ | model.labelLike+ ),
            ( model.noteLike | model.biblLike )*
         )?
      ),
      population*
   )
}
```

---

**<pos>** (part of speech) indicates the part of speech assigned to a dictionary headword such as noun, verb, or adjective. [*10.3.2. Grammatical Information*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    - *@datcat*
    - *@valueDatcat*
    - *@targetDatcat*
  - att.lexicographic.normalized
    - *@norm*
    - *@orig*

*Member of* model.entryPart model.lexicalRefinement

*Contained by*

*core:* cit

*dictionaries:* dictScrap entryFree form gramGrp

*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<entry>
  <form>
    <orth>isotope</orth>
  </form>
  <gramGrp>
    <pos>adj</pos>
  </gramGrp>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element pos
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   macro.paraContent}
```

---

## <post>

a written (or spoken) contribution to a CMC interaction which has been composed (or recorded) by its author in its entirety before being transmitted over a network (e.g., the internet) and made available on the monitor or screen of the other parties en bloc. [*9.3.1. CMC Posts*]

*Module* cmc — Computer-mediated Communication

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.ascribed
  - *@who*
- att.canonical
  - *@key*
  - *@ref*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*

  * *@notAfter-iso*
  * *@from-iso*
  * *@to-iso*
  - att.datable.w3c
  * *@when*
  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*
- att.docStatus
  - *@status*
- att.fragmentable
  - *@part*
- att.indentation
  - *@indentLevel*
- att.timed
  - *@start*
  - *@end*

  - att.duration
  * att.duration.iso
  · *@dur-iso*
  * att.duration.w3c
  · *@dur*
- att.typed
  - *@type*
  - *@subtype*

`@modality` written or spoken mode.

*Status* Recommended

*Datatype* teidata.enumerated

*Suggested values include:* **written** for stretches of text and/or multimodal elements

**spoken** (for audio (or audio-visual) posts)

*Note* This attribute is intended to convey information about the CMC environment in which the post was created, as opposed to the MIME type of the content conveyed, which is indicated by the elements contained within the <post>.

`@replyTo` indicates to which previous post the current post replies or refers.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Note* This attribute should be used to encode 'technical' reply information that is due to a formal reply action (such as activating a 'reply' button in the client software) and that is formally represented in the source, e.g. in the 'references' field of a Usenet message header or in the subject line of a forum post. It should not be used for inferred or interpreted reply relations based only on ambiguous signs such as linguistic discourse markers or indentation.

*Member of* model.cmc

*Contained by*

*drama:* castList epilogue performance prologue set

*figures:* figure

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic head hi index l label lb lg list listBibl measure measureGrp media meeting mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Even though in some CMC genres (e.g., WhatsApp), the other parties are informed by an automated alert about the fact that another party is currently composing a new post, they cannot track the process of verbalization, i.e. how the written utterance emerges in the entry form on the user interface of its author. It is not until the author performs a 'posting' action (e.g., by hitting the 'enter' key or by activating a 'send' button with the mouse) that the result of the composition process — the post — is made available for the other parties. From the perspective of its addressees/readers, a post is a stretch of text that has been composed in advance. Different from spoken utterances but similar to edited text, (1) the reader of a post does not have access to the verbalization process and (2) the mental processing of a post is not possible until the process of verbalization (the composition) has been finished and the posting action has been performed by the author. Posts occur in a broad range of written CMC genres: as user 'messages' in chats and WhatsApp dialogues, as SMS messages, as tweets in Twitter timelines, as individual comments following a status update on Facebook pages, as posts in forum threads, as contributions on Wikipedia talk pages or in the comments section of a weblog. [Element adopted from the DeRiK schema. Aspects of the model adopted from the CoMeRe Schema]. The <post> element is a member of the model class model.common (via model.CMC). This allows for using (and combining) occurrences of <post>, <u>,

<incident>, and <kinesic>in one and the same  in order to be able to represent the combined use of written and spoken utterances and nonverbal acts on the GUI of multimodal CMC environments.

The use of the *modality* attribute is recommended, but not required. Its value is either written or spoken. Because spoken language differs fundamentally from written language, when the value is spoken projects may wish to limit the content of the <post> to character data and only those elements available in the <u> element, thus avoiding elements like <p> or <head> that are used for the encoding of written text.

*Example* standard user postings in chat

```
<post type="standard" generatedBy="human"
  who="#A05" rend="color:red">
  <p>oh ja, in die Wärme...</p>
</post>
<post type="standard" generatedBy="human"
  who="#A03" rend="color:red">
  <p>nein, hast du nicht !!</p>
</post>
```

*Example* chat message-type ”system”: user enters a chatroom, ”system” defined in listPerson

```
<listPerson>
  <person xml:id="system">
    <persName>system</persName>
  </person>
<!-- ... -->
</listPerson>
<!-- ... -->
<post type="event" generatedBy="system"
  who="#system" rend="color:navy">
  <p>
    <name type="nickname" corresp="#A07">Interseb</name> betritt den Raum.</p>
</post>
```

*Example* action message in chat

```
<post type="event" generatedBy="human"
  who="#A02" rend="color:blue">
  <p>
    <name type="nickname" corresp="#A02">McMike</name> hechtet zum Steuer </p>
</post>
```

*Example* chat messages incl. interaction words (types: inflective and acronym) and an interaction phrase

```
<post type="standard" generatedBy="human"
  who="#A03" rend="color:green">
  <p>
    <w type="AKWIND">*</w>
    <w type="AKW" subtype="inflective">schwank</w>
    <w type="AKWIND">*</w>
<!-- ... -->
    <w type="AKWIND">*</w>
    <w type="AKW">seekrank</w>
    <w type="AKWIND">*</w>
  </p>
</post>
<!-- ... -->
<post type="standard" generatedBy="human"
```

```
  who="#A02"  rend="color:blue">
  <p>
    <w  type="AKWIND">*</w>
    <w  type="AKW"  subtype="acronym">lol</w>
    <w  type="AKWIND">*</w> ... </p>
</post>
<!-- ... -->
<post  type="standard"  generatedBy="human"
  who="#A01"  rend="color:black">
  <p>
    <phr  type="AKW">
      <w  type="AKWIND">*</w>
      <name  type="nickname"  corresp="#A02">McMike</name> einen fallschirm
<w  type="AKW"  subtype="inflective">umbind</w>
      <w  type="AKWIND">*</w>
    </phr>
  </p>
</post>
```

*Example* user posting incl. an emoticon

```
<post  type="standard"  generatedBy="human"
  who="#A01"  rend="color:black">
  <p> naja, der war gestrandet <w  type="emoticon"  subtype="ascii">:(</w>
  </p>
</post>
```

*Example* Twitter: Tweet containing an addressing term and a hashtag expression

```
<post  type="standard"  generatedBy="human"
  who="#A01"  when="2015-11-18">
  <p>
    <ref  corresp="#A02"
      target="https://twitter.com/textarchiv">
      <w  type="ADRIND">@</w>
      <w  type="ADR">textarchiv</w>
    </ref> Erfurter
     Europäische Welt- und Staats-Geschichte: 20 Ausgaben von 1744 aus dem MKHZ
     bereits in DTAQ
<ref  target="http://www.deutschestextarchiv.de/dtaq/about">
http://www.deutschestextarchiv.de/dtaq/about</ref>
    <ref  target="https://twitter.com/hashtag/dtaclarin14?src=hash">
      <w  type="hashtag">#dtaclarin14</w>
    </ref>
  </p>
</post>
```

*Example* example from Wikipedia discussion

```
<post  type="standard"  generatedBy="human"
  indentLevel="3"  synch="#t02622878"  who="#WU00018921">
  <p> Wie du siehst hab ich die Lemma geändert, danke für den Hinweis, ich war
nämlich
     selbst auch etwas unsicher bei der ganzen Sache und bin jetzt damit auch
viel
     glücklicher!--</p>
  <signed  rend="inline"
```

```
    generatedBy="template">
    <gap  reason="signatureContent"/>
    <time  generatedBy="template">12:01, 12. Jun. 2009 (CEST)</time>
  </signed>
</post>
```

*Example* Blog comment: Die Anthropodizee-Frage. Wer den Himmel leerräumt, schafft die Menschheit ab. by Volker Birk.
??ttps://scilogs.spektrum.de/natur-des-glaubens/die-anthropodizee-frage-wer-himmel/#comment-31899

```
<post  xml:id="p5"  type="comment"  who="#u4"
  synch="#t005"  replyTo="#p4">
  <p>“Wenn Sie diesen Gruppen also “mangelnde Bildung“ attestieren wollen, so
     verwenden Sie bereits einen bestimmten, kulturgebundenen
Bildungsbegriff.”</p>
  <p>Ich hoffe doch, wir können beim Bildungsbegriff der Aufklärung bleiben.
Wer das
    nicht möchte, hat die Wissenschaft verlassen.</p>
</post>
```

*Example* Usenet reply that comments on material quoted in an earlier message.

```
<post  when="1990-10-11"  who="#dmerritt"
  xml:id="sgml-tei-3"  replyTo="#sgml-tei-1">
  <time  generatedBy="system">04:37:51 UTC</time>
  <p>
    <seg  generatedBy="system">In article
    <ref  target="#sgml-tei-1"><2****.**********0@m*******.***.*****.**c.uk></ref>
S.*.*.*****z@e**.*****.**c.uk (Sebastian Rahtz) writes:</seg>
    <cit>
      <bibl>[for] Lou Burnard:</bibl>
      <quote>the Text Encoding Initiative document is a many-splendoured thing,
            redolent of the mysterious east, caribbean evenings and the scent of
            fresh pine in the himalayas [...]</quote>
    </cit>
  </p>
  <p>Very poetic.</p>
  <p>But really, what is it?</p>
  <signed  generatedBy="human">Doug</signed>
  <signed  generatedBy="template">Doug Merritt ****@e***.********y.edu
(ucbvax!eris!doug) or uunet.uu.net!crossck!dougm
  </signed>
</post>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.headLike"/>
  <elementRef key="opener"/>
  <classRef key="model.pLike"/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.global"/>
  <elementRef key="lg"/>
```

```
  <classRef key="model.lLike"/>
  <classRef key="model.divBottom"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element post
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.attributes,
   att.canonical.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.docStatus.attributes,
   att.fragmentable.attributes,
   att.indentation.attributes,
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   att.typed.attributes,
   attribute modality { "written" | "spoken" }?,
   attribute replyTo { list { + } }?,
   (
      text
    | model.headLike    | opener    | model.pLike    | model.gLike    | model.phrase    |
}
```

---

## <postBox>

(postal box or post office box) contains a number or other identifier for some postal delivery point other than a street address. [*3.6.2. Addresses*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change

    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.addrPart

*Contained by*

*core:* address

*May contain* Character data only

*Note* The position and nature of postal codes is highly country-specific; the conventions appropriate to the country concerned should be used.

*Example*

```
<postBox>P.O. Box 280</postBox>
```

*Example*

```
<postBox>Postbus 532</postBox>
```

*Content model* `<content> <textNode/></content>`

*Schema Declaration*

```
element postBox
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
```

```
   text
}
```

---

**<postCode>** (postal code) contains a numerical or alphanumeric code used as part of a postal address to simplify sorting or delivery of mail. [*3.6.2. Addresses*]

*Module* core — Elements Available in All TEI Documents

*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  - * *@ana*
- – att.global.change
  - * *@change*
- – att.global.facs
  - * *@facs*
- – att.global.linking
  - * *@corresp*
  - * *@synch*
  - * *@sameAs*
  - * *@copyOf*
  - * *@next*
  - * *@prev*
  - * *@exclude*
  - * *@select*
- – att.global.rendition
  - * *@rend*
  - * *@style*
  - * *@rendition*
- – att.global.responsibility
  - * *@cert*
  - * *@resp*
- – att.global.source
  - * *@source*

*Member of* model.addrPart

*Contained by*

*core:* address

*May contain* Character data only

*Note* The position and nature of postal codes is highly country-specific; the conventions appropriate to the country concerned should be used.

*Example*

```
<postCode>HR1 3LR</postCode>
```

*Example*

```
<postCode>60142-7</postCode>
```

*Content model* `<content> <textNode/></content>`

*Schema Declaration*

```
element postCode
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   text
}
```

---

## <postscript>

contains a postscript, e.g. to a letter. [*4.2. Elements Common to All Divisions*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.placement
  - *@place*
- att.written
  - *@hand*

*Member of* model.divBottomPart
*Contained by*
*cmc:* post
*core:* lg list
*drama:* epilogue performance prologue
*figures:* figure table
*textstructure:* back body div div1 div2 div3 div4 div5 div6 div7 front group postscript
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*cmc:* post
*core:* bibl biblStruct cb cit desc ellipsis gap gb head index l label lb lg list listBibl milestone note noteGrp p pb q quote said sp stage
*dictionaries:* entry entryFree superEntry
*drama:* camera caption castList move sound spGrp tech view
*figures:* figure notatedMusic table
*header:* biblFull
*iso-fs:* fLib fs fvLib
*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* msDesc
*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation
*nets:* eTree forest graph listForest tree
*spoken:* annotationBlock incident kinesic pause shift u vocal writing
*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef
*textcrit:* app listApp listWit witDetail
*textstructure:* closer floatingText opener postscript signed trailer
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
*Example*

```
<div  type="letter">
  <opener>
    <dateline>
      <placeName>Rimaone</placeName>
      <date  when="2006-11-21">21 Nov 06</date>
    </dateline>
    <salute>Dear Susan,</salute>
  </opener>
  <p>Thank you very much for the assistance splitting those
     logs. I'm sorry about the misunderstanding as to the size of
     the task. I really was not asking for help, only to borrow the
     axe. Hope you had fun in any case.</p>
  <closer>
    <salute>Sincerely yours,</salute>
    <signed>Seymour</signed>
  </closer>
  <postscript>
    <label>P.S.</label>
    <p>The collision occurred on <date  when="2001-07-06">06 Jul 01</date>.</p>
  </postscript>
</div>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.global"/>
   <classRef key="model.divTopPart"/>
  </alternate>
  <classRef key="model.common"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.global"/>
   <classRef key="model.common"/>
  </alternate>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divBottomPart"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element postscript
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.cmc.attributes,
   att.placement.attributes,
   att.written.attributes,
   (
      ( model.global | model.divTopPart )*,
      model.common,
      ( model.global | model.common )*,
      ( ( model.divBottomPart, model.global* )* )
   )
}
```

---

**<precision>** indicates the numerical accuracy or precision associated with some aspect of the text markup. [*22.2. Indications of Precision*]

*Module* certainty — Certainty, Precision, and Responsibility

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

  * *@source*

- att.datable
  - *@period*

  - att.datable.custom
  * *@when-custom*
  * *@notBefore-custom*
  * *@notAfter-custom*
  * *@from-custom*
  * *@to-custom*
  * *@datingPoint*
  * *@datingMethod*
  - att.datable.iso
  * *@when-iso*
  * *@notBefore-iso*
  * *@notAfter-iso*
  * *@from-iso*
  * *@to-iso*
  - att.datable.w3c
  * *@when*
  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*

- att.ranging
  - *@atLeast*
  - *@atMost*
  - *@min*
  - *@max*
  - *@confidence*

- att.scoping
  - *@target*
  - *@match*

`@precision` characterizes the precision of the element or attribute pointed to by the <precision> element.
*Status* Optional
*Datatype* teidata.certainty

`@stdDeviation` supplies a standard deviation associated with the value in question.
*Status* Optional
*Datatype* teidata.numeric

*Member of* model.certLike model.global.meta
*Contained by*
*analysis:* cl interp m phr s span w

*certainty:* certainty precision respons

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gap gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab join seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc climate country death district education eventName faith floruit forename genName gender geogFeat geogName langKnowledge langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName population region residence roleName settlement sex socecStatus state surname terrain trait

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc space substJoin supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*certainty:* certainty precision respons

*core:* desc

*Example*

```
<date xml:id="date001" notBefore="0014"
  notAfter="0064">Within about 50
years of the death of Augustus</date>
<precision target="#date001"
  match="@notBefore" precision="high"/>
<precision target="#date001"
  match="@notAfter" precision="low"/>
```

*Example*

```
<date notBefore="1632" notAfter="1642">Not more than ten years before the
start of the Civil War
<precision match="@notBefore"
```

```
        precision="low"/>
      <precision  match="@notAfter"
        precision="high"/>
    </date>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.descLike"/>
  <classRef key="model.certLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element precision
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.ranging.attributes,
   att.scoping.attributes,
   attribute precision { text }?,
   attribute stdDeviation { text }?,
   ( model.descLike | model.certLike )*
}
```

---

## <pRef>

(pronunciation reference) in a dictionary example, indicates a reference to the pronunciation(s) of the headword. [*10.4. Headword and Pronunciation References*]

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    - *@datcat*
    - *@valueDatcat*
    - *@targetDatcat*
  - att.lexicographic.normalized
    - *@norm*
    - *@orig*
- att.notated
  - *@notation*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*

*Member of* model.ptrLike.form

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth pRef per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*dictionaries:* pRef

*gaiji:* g

character data

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <elementRef key="pRef"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element pRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
```

```
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.notated.attributes,
   att.pointing.attributes,
   ( text | model.gLike | pRef )*
}
```

---

**<prefixDef>** (prefix definition) defines a prefixing scheme used in `teidata.pointer` values, showing how abbreviated URIs using the scheme may be expanded into full URIs. [*17.2.3. Using Abbreviated Pointers*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

  * *@source*

- att.patternReplacement
  - *@matchPattern*
  - *@replacementPattern*

`@ident` supplies a name which functions as the prefix for an abbreviated pointing scheme such as a private URI scheme. The prefix constitutes the text preceding the first colon.
*Status* Required
*Datatype* teidata.prefix
*Note* The value is limited to `teidata.prefix` so that it may be mapped directly to a URI prefix.

*Contained by:* listPrefixDef
*May contain*
*core:* p
*linking:* ab
*Note* The abbreviated pointer may be dereferenced to produce either an absolute or a relative URI reference. In the latter case it is combined with the value of *xml:base* in force at the place where the pointing attribute occurs to form an absolute URI in the usual manner as prescribed by XML Base.
*Example*

```
<prefixDef  ident="ref"
  matchPattern="([a-z]+)"
  replacementPattern="../../references/references.xml#$1">
  <p> In the context of this project, private URIs with
     the prefix "ref" point to <gi>div</gi> elements in
     the project's global references.xml file.
  </p>
</prefixDef>
```

*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="0"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element prefixDef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.patternReplacement.attributes,
   attribute ident { text },
   model.pLike*
}
```

**<preparedness>** (preparedness) describes the extent to which a text may be regarded as prepared or spontaneous. [*16.2.1. The Text Description*]

*Module* corpus — Language Corpora

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - ~~type~~
  - @subtype

`@type` a keyword characterizing the type of preparedness.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **none** spontaneous or unprepared

**scripted** follows a script
**formulaic** follows a predefined set of conventions
**revised** polished or revised before presentation

*Member of* model.textDescPart
*Contained by*
*corpus:* textDesc
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit
*dictionaries:* lang
*figures:* figure notatedMusic
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef tag val
*textcrit:* app witDetail
*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data
*Example*

```
<preparedness type="none"/>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element preparedness
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.phraseSeq.limited}
```

**<principal>** (principal researcher) supplies the name of the principal researcher responsible for the creation of an electronic text. [*2.2.1. The Title Statement*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*

    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*

*Member of* model.respLike

*Contained by*

*core:* bibl

*header:* editionStmt titleStmt

*msdescription:* msItem

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin

character data

*Example*

```
<principal  ref="http://viaf.org/viaf/105517912">Gary Taylor</principal>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element principal
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   macro.phraseSeq.limited}
```

---

## <profileDesc>

<profileDesc> (text-profile description) provides a detailed description of non-bibliographic aspects of a text, specifically the languages and sublanguages used, the situation in which it was produced, the participants and their setting. [*2.4. The Profile Description 2.1.1. The TEI Header and Its Components*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.teiHeaderPart

*Contained by*

*header:* biblFull teiHeader

*May contain*

*corpus:* particDesc settingDesc textDesc

*header:* abstract calendarDesc correspDesc creation langUsage textClass

*transcr:* handNotes listTranspose

*Note* Although the content model permits it, it is rarely meaningful to supply multiple occurrences for any of the child elements of <profileDesc> unless these are documenting multiple texts.

*Example*

```
<profileDesc>
  <langUsage>
    <language ident="fr">French</language>
  </langUsage>
  <textDesc n="novel">
    <channel mode="w">print; part issues</channel>
    <constitution type="single"/>
    <derivation type="original"/>
    <domain type="art"/>
    <factuality type="fiction"/>
    <interaction type="none"/>
    <preparedness type="prepared"/>
    <purpose type="entertain" degree="high"/>
    <purpose type="inform" degree="medium"/>
  </textDesc>
  <settingDesc>
    <setting>
      <name>Paris, France</name>
      <time>Late 19th century</time>
    </setting>
  </settingDesc>
</profileDesc>
```

*Content model*

```
<content>
```

```
 <classRef key="model.profileDescPart"
  minOccurs="0" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element profileDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   model.profileDescPart*
}
```

---

**<projectDesc>** (project description) describes in detail the aim or purpose for which an electronic file was encoded, together with any other relevant information concerning the process by which it was assembled or collected. [*2.3.1. The Project Description 2.3. The Encoding Description 16.3.2. Declarable Elements*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.declarable
  - *@default*

*Member of* model.encodingDescPart

*Contained by*

*header:* encodingDesc

*May contain*

*core:* p

*linking:* ab

*Example*

```
<projectDesc>
  <p>Texts collected for use in the Claremont Shakespeare Clinic, June 1990</p>
</projectDesc>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:projectDesc"/> </sch:pattern>

*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element projectDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   model.pLike+
}
```

---

**<prologue>** (prologue) contains the prologue to a drama, typically spoken by an actor out of character, possibly in association with a particular performance or venue. [*7.1.2. Prologues and Epilogues 7.1. Front and Back Matter* ]

*Module* drama — Performance Texts
*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  * *@ana*
- – att.global.change
  * *@change*
- – att.global.facs
  * *@facs*
- – att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- – att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- – att.global.responsibility
  * *@cert*
  * *@resp*
- – att.global.source
  * *@source*

*Member of* model.frontPart.drama
*Contained by*
*textcrit:* lem rdg
*textstructure:* back front
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*cmc:* post
*core:* bibl biblStruct cb cit desc ellipsis gap gb head index l label lb lg list listBibl meeting milestone note noteGrp p pb q quote said sp stage
*dictionaries:* entry entryFree superEntry
*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* argument byline closer dateline docAuthor docDate epigraph floatingText opener postscript salute signed trailer

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Example*

```
<prologue>
  <sp>
    <l>Wits, like physicians never can agree,</l>
    <l>When of a different society.</l>
    <l>New plays are stuffed with wits, and with deboches,</l>
    <l>That crowd and sweat like cits in May-Day coaches.</l>
  </sp>
  <trailer>Written by a person of quality</trailer>
</prologue>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divTop"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.common"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divBottom"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element prologue
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      ( model.divTop | model.global )*,
      ( ( model.common, model.global* )+ ),
      ( ( model.divBottom, model.global* )* )
   )
}
```

---

**<pron>** (pronunciation) contains the pronunciation(s) of the word. [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*

    * *@resp*
  - att.global.source
    * *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*
  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.lexicographic.normalized
    * *@norm*
    * *@orig*
- att.notated
  - *@notation*
- att.partials
  - *@extent*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.entryPart model.formPart

*Contained by*

*core:* cit

*dictionaries:* dictScrap entryFree form

*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The values used to specify the notation may be taken from any appropriate project-defined list of values. Typical values might be *IPA*, *Murray*, for example.

*Example*

```
<entry>
  <form>
    <orth>obverse</orth>
    <pron>'äb-`ərs</pron>,
  <pron  extent="pref">äb-`</pron>, <pron  extent="pref">əb-`</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
</entry>
```

*Example*

```
<entry>
  <form>
    <orth>transcription</orth>
    <pron  notation="IPA">trænskrɪpʃən</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element pron
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
```

```
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.notated.attributes,
   att.partials.attributes,
   att.typed.attributes,
   macro.paraContent}
```

---

**<provenance>** (provenance) contains any descriptive or other information concerning a single identifiable episode during the history of a manuscript, manuscript part, or other object after its creation but before its acquisition. [*11.8. History*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*

    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.typed
  - *@type*
  - *@subtype*

*Contained by:* *msDescription:* history
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound spGrp tech view
*figures:* figure formula notatedMusic table
*gaiji:* g
*header:* biblFull idno
*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<provenance>Listed as the property of Lawrence Sterne in 1788.</provenance>
<provenance>Sold at Sothebys in 1899.</provenance>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element provenance
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.typed.attributes,
   macro.specialPara}
```

---

## <ptr>

(pointer) defines a pointer to another location. [*3.7. Simple Links and Cross-References 17.1. Links*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*

  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cReferencing
  - *@cRef*
- att.cmc
  - *@generatedBy*
- att.declaring
  - *@decls*
- att.internetMedia
  - *@mimeType*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*
- att.typed
  - *@type*

– *@subtype*

*Member of* model.annotationPart.body model.ptrLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine analytic author bibl biblScope biblStruct cit citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned monogr name note num orig p pubPlace publisher q quote rb ref reg relatedItem resp rs rt said series sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc notatedMusic

*header:* application authority catDesc change classCode correspContext creation distributor edition extent funder geoDecl handNote language licence principal publicationStmt rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab altGrp annotation joinGrp linkGrp seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education event eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person place placeName region residence roleName settlement sex socecStatus surname

*nets:* eLeaf eTree

*spoken:* transcriptionDesc u writing

*tagdocs:* eg listRef valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus transpose

*verse:* metSym rhyme

*May contain* Empty element

*Note* The *target* and *cRef* attributes are mutually exclusive.

*Example*

```
<ptr target="#p143 #p144"/>
<ptr target="http://www.tei-c.org"/>
<ptr cRef="1.3.4"/>
```

*Schematron* <sch:rule context="tei:ptr"> <sch:report test="@target and @cRef">Only one of the attributes @target and @cRef may be supplied on <sch:name/>.</sch:report> </sch:rule>

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element ptr
{
```

```
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cReferencing.attributes,
   att.cmc.attributes,
   att.declaring.attributes,
   att.internetMedia.attributes,
   att.pointing.attributes,
   att.typed.attributes,
   empty
}
```

---

**<publicationStmt>** (publication statement) groups information concerning the publication or distribution of an electronic or other text. [*2.2.4. Publication, Distribution, Licensing, etc. 2.2. The File Description*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*

- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Contained by:* biblFull fileDesc

*May contain*

*core:* address date p ptr pubPlace publisher ref

*header:* authority availability distributor idno

*linking:* ab

*tagdocs:* listRef

*Note* Where a publication statement contains several members of the `model.publicationStmtPart.agency` or `model.publicationStmtPart.detail` classes rather than one or more paragraphs or anonymous blocks, care should be taken to ensure that the repeated elements are presented in a meaningful order. It is a conformance requirement that elements supplying information about publication place, address, identifier, availability, and date be given following the name of the publisher, distributor, or authority concerned, and preferably in that order.

*Example*

```
<publicationStmt>
  <publisher>C. Muquardt </publisher>
  <pubPlace>Bruxelles & Leipzig</pubPlace>
  <date  when="1846"/>
</publicationStmt>
```

*Example*

```
<publicationStmt>
  <publisher>Chadwyck Healey</publisher>
  <pubPlace>Cambridge</pubPlace>
  <availability>
    <p>Available under licence only</p>
  </availability>
  <date  when="1992">1992</date>
</publicationStmt>
```

*Example*

```
<publicationStmt>
  <publisher>Zea Books</publisher>
  <pubPlace>Lincoln, NE</pubPlace>
  <date>2017</date>
  <availability>
    <p>This is an open access work licensed under a Creative Commons
Attribution 4.0 International license.</p>
  </availability>
  <ptr  target="http://digitalcommons.unl.edu/zeabook/55"/>
</publicationStmt>
```

*Content model*

```
<content>
 <alternate>
```

```
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.publicationStmtPart.agency"/>
   <classRef key="model.publicationStmtPart.detail"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element publicationStmt
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      (
         (
            model.publicationStmtPart.agency,
            model.publicationStmtPart.detail*
         )+
      )
    | model.pLike+
   )
}
```

---

**<publisher>** (publisher) provides the name of the organization responsible for the publication or distribution of a bibliographic item. [*3.12.2.4. Imprint, Size of a Document, and Reprint Information 2.2.4. Publication, Distribution, Licensing, etc.*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs

    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.canonical
  - *@key*
  - *@ref*

*Member of* model.imprintPart model.publicationStmtPart.agency

*Contained by*

*core:* bibl imprint

*header:* publicationStmt

*textstructure:* docImprint

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Use the full form of the name by which a company is usually referred to, rather than any abbreviation of it which may appear on a title page

*Example*

```
<imprint>
  <pubPlace>Oxford</pubPlace>
  <publisher>Clarendon Press</publisher>
  <date>1987</date>
</imprint>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element publisher
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   macro.phraseSeq}
```

---

## <pubPlace> (publication place) contains the name of the place where a bibliographic item was published.

[*3.12.2.4. Imprint, Size of a Document, and Reprint Information*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
    - *@key*
    - *@ref*

*Member of* model.imprintPart model.publicationStmtPart.detail

*Contained by*

*core:* bibl imprint

*header:* publicationStmt

*textstructure:* docImprint

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<publicationStmt>
  <publisher>Oxford University Press</publisher>
  <pubPlace>Oxford</pubPlace>
  <date>1989</date>
</publicationStmt>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element pubPlace
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   macro.phraseSeq}
```

---

**<punctuation>** specifies editorial practice adopted with respect to punctuation marks in the original. [*2.3.3. The Editorial Practices Declaration 3.2. Treatment of Punctuation*]

*Module* header — The TEI Header

*Attributes* • att.global

- *@xml:id*
- *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.declarable
  - *@default*

`@marks` indicates whether or not punctation marks have been retained as content within the text.
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **none** no punctuation marks have been retained
**some** some punctuation marks have been retained
**all** all punctuation marks have been retained

`@placement` indicates the positioning of punctuation marks that are associated with marked up text as being encoded within the element surrounding the text or immediately before or after it.
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **internal** punctuation marks found at the start or end of a marked up text component are included within its surrounding element;
**external** punctuation marks found at the start or end of a marked up text component appear immediately before or after the surrounding element

*Member of* model.editorialDeclPart
*Contained by*
*header:* editorialDecl
*May contain*
*core:* p
*linking:* ab
*Example*

```
<punctuation  marks="all"
  placement="internal">
  <p>All punctuation marks in the source text have been retained and
represented using the
    appropriate Unicode code point. In cases where a punctuation mark and
nearby markup convey
    the same information (for example, a sentence ends with a question mark
and is also tagged
    as <gi>s</gi>) the punctuation mark is captured as content within the
element.</p>
</punctuation>
```

*Example* External placement of punctuation:

```
<p>I would agree with Saint Augustine that “<quote>An unjust law is no law at
all</quote>.”</p>
```

*Example* Internal placement of punctuation:

```
<p>I would agree with Saint Augustine that <quote>“An unjust law is no law at
all.”</quote>
</p>
```

*Schematron* <sch:pattern is-a=”declarable”> <sch:param name=”tde” value=”tei:punctuation”/> </sch:pattern>
*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="0"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element punctuation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   attribute marks { "none" | "some" | "all" }?,
   attribute placement { "internal" | "external" }?,
```

```
    model.pLike*
}
```

---

**<purpose>** characterizes a single purpose or communicative function of the text. [*16.2.1. The Text Description*]

*Module* corpus — Language Corpora

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - ~~type~~
  - @subtype

`@type` specifies a particular kind of purpose.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **persuade** didactic, advertising, propaganda, etc.

**express** self expression, confessional, etc.

**inform** convey information, educate, etc.

**entertain** amuse, entertain, etc.

`@degree` specifies the extent to which this purpose predominates.

*Status* Optional

*Datatype* teidata.certainty

*Note* Values should be interpreted as follows.

**high** this purpose is predominant

**medium** this purpose is intermediate

**low** this purpose is weak

**unknown** extent unknown

*Contained by:* textDesc

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin

character data

*Note* Usually empty, unless some further clarification of the type attribute is needed, in which case it may contain running prose

*Example*

```
<purpose type="persuade" degree="high"/>
<purpose type="entertain" degree="low"/>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element purpose
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { "persuade" | "express" | "inform" | "entertain" }?,
   attribute degree { text }?,
   macro.phraseSeq.limited}
```

---

**<q>** (quoted) contains material which is distinguished from the surrounding text using quotation marks or a similar method, for any one of a variety of reasons including, but not limited to: direct speech or thought, technical terms or jargon, authorial distance, quotations from elsewhere, and passages that are mentioned but not used. [*3.3.3. Quotation*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*

  * *@rendition*
  - att.global.responsibility
  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.ascribed.directed
  - *@toWhom*
  - att.ascribed
  * *@who*
- att.cmc
  - *@generatedBy*

`@type` (type) may be used to indicate whether the offset passage is spoken or thought, or to characterize it more finely.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **spoken** (spoken) representation of speech
**thought** (thought) representation of thought, e.g. internal monologue
**written** (written) quotation from a written source
**soCalled** (so called) authorial distance
**foreign** (foreign) foreign words
**distinct** (distinct) linguistically distinct
**term** technical term
**emph** (emph) rhetorically emphasized
**mentioned** (mentioned) refering to itself, not its normal referent

*Member of* model.common model.hiLike
*Contained by*
*analysis:* cl m phr s span w
*cmc:* post
*core:* abbr add addrLine author bibl biblScope cit citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled sp speaker stage street term textLang time title unclear unit
*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castItem castList epilogue performance prologue role roleDesc set sound tech view
*figures:* cell figDesc figure formula
*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage
*iso-fs:* fDescr fsDescr
*linking:* ab seg
*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation

objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* argument body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint epigraph imprimatur opener postscript salute signed titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May be used to indicate that a passage is distinguished from the surrounding text for reasons concerning which no claim is made. When used in this manner, <q> may be thought of as syntactic sugar for <hi> with a value of *rend* that indicates the use of such mechanisms as quotation marks.

*Example*

```
It is spelled <q>Tübingen</q> — to enter the
letter <q>u</q> with an umlaut hold down the <q>option</q> key and press
<q>0 0 f c</q>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element q
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.directed.attributes,
   att.ascribed.attributes,
   att.cmc.attributes,
   attribute type
   {
      "spoken"
    | "thought"
    | "written"
    | "soCalled"
    | "foreign"
    | "distinct"
    | "term"
    | "emph"
    | "mentioned"
   }?,
   macro.specialPara}
```

---

**<quotation>** (quotation) specifies editorial practice adopted with respect to quotation marks in the original.
[*2.3.3. The Editorial Practices Declaration 16.3.2. Declarable Elements*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable
  - *@default*

`@marks` (quotation marks) indicates whether or not quotation marks have been retained as content within the text.
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **none** no quotation marks have been retained
**some** some quotation marks have been retained
**all** all quotation marks have been retained

*Member of* model.editorialDeclPart
*Contained by*
*header:* editorialDecl
*May contain*
*core:* p
*linking:* ab
*Example*

```
<quotation  marks="none">
  <p>No quotation marks have been retained. Instead, the <att>rend</att>
attribute on the
  <gi>q</gi> element is used to specify what kinds of quotation mark was used,
according
    to the following list: <list  type="gloss">
```

```
      <label>dq</label>
      <item>double quotes, open and close</item>
      <label>sq</label>
      <item>single quotes, open and close</item>
      <label>dash</label>
      <item>long dash open, no close</item>
      <label>dg</label>
      <item>double guillemets, open and close</item>
    </list>
  </p>
</quotation>
```

*Example*

```
<quotation  marks="all">
  <p>All quotation marks are retained in the text and are represented by
appropriate Unicode
    characters.</p>
</quotation>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:quotation"/> </sch:pattern>

*Schematron* <sch:rule context="tei:quotation"> <sch:report test="not( @marks ) and not( tei:p )"> On <sch:name/>, either the @marks attribute should be used, or a paragraph of description provided </sch:report> </sch:rule>

*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="0"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element quotation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   attribute marks { "none" | "some" | "all" }?,
   model.pLike*
}
```

---

**<quote>** (quotation) contains a phrase or passage attributed by the narrator or author to some agency external to the text. [*3.3.3. Quotation 4.3.1. Grouped Texts*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*

  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.msExcerpt
  - *@defective*
- att.notated
  - *@notation*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.biblPart model.quoteLike

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author bibl biblScope cit citedRange corr del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p

pubPlace publisher q quote rb ref reg rs rt said sic soCalled sp speaker stage street term textLang title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* actor camera caption castList epilogue performance prologue role roleDesc set sound tech view

*figures:* cell figDesc figure

*header:* change distributor edition extent geoDecl handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material msItem musicNotation objectType origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* If a bibliographic citation is supplied for the source of a quotation, the two may be grouped using the <cit> element.

*Example*

```
Lexicography has shown little sign of being affected by the
work of followers of J.R. Firth, probably best summarized in his
slogan, <quote>You shall know a word by the company it
keeps</quote>
<ref>(Firth, 1957)</ref>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element quote
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.msExcerpt.attributes,
   att.notated.attributes,
   att.typed.attributes,
   macro.specialPara}
```

---

## <rb>

(ruby base) contains the base text annotated by a ruby gloss. [*3.4.2. Ruby Annotations*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
- att.global.analytic
  - *@ana*

  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.typed
  - *@type*
  - *@subtype*

*Contained by* ruby

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example* The word 你好 *nǐ hǎo* (hello) is glossed in pinyin to provide a pronunciation guide.

```
<p  xml:lang="zh">
<!--…-->
  <ruby>
    <rb>你</rb>
    <rt  place="above">nǐ</rt>
  </ruby>
  <ruby>
    <rb>好</rb>
    <rt  place="above">hǎo</rt>
  </ruby>
<!--…-->
</p>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element rb
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

## <rdg>

(reading) contains a single reading within a textual variation. [*13.1. The Apparatus Entry, Readings, and Witnesses*]

*Module* textcrit — Critical Apparatus

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.textCritical
  - *@type*
  - *@cause*
  - *@varSeq*
  - *@require*

- att.typed
  - ~~type~~
  - @subtype

- att.written
  - *@hand*

- att.witnessed
  - *@wit*

*Member of* model.rdgLike
*Contained by*
*textcrit:* app rdgGrp

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList epilogue move performance prologue set sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app lacunaEnd lacunaStart listApp listWit wit witDetail witEnd witStart

*textstructure:* argument byline div docAuthor docDate docEdition docImprint docTitle epigraph floatingText imprimatur titlePage titlePart

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<rdg wit="#Ra2">Eryment</rdg>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.divLike"/>
  <classRef key="model.divPart"/>
  <elementRef key="titlePage"/>
  <elementRef key="argument"/>
  <elementRef key="byline"/>
  <elementRef key="docAuthor"/>
  <elementRef key="docDate"/>
  <elementRef key="docEdition"/>
  <elementRef key="docImprint"/>
  <elementRef key="docTitle"/>
  <elementRef key="epigraph"/>
```

```
  <elementRef key="imprimatur"/>
  <elementRef key="titlePart"/>
  <elementRef key="epilogue"/>
  <elementRef key="performance"/>
  <elementRef key="prologue"/>
  <elementRef key="set"/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.global"/>
  <classRef key="model.rdgPart"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element rdg
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.textCritical.attributes,
   att.written.attributes,
   att.witnessed.attributes,
   (
      text
    | model.divLike    | model.divPart    | titlePage    | argument    | byline    | docAu
}
```

---

**<rdgGrp>** (reading group) within a textual variation, groups two or more readings perceived to have a genetic relationship or other affinity. [*13.1. The Apparatus Entry, Readings, and Witnesses*]

*Module* textcrit — Critical Apparatus

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*

  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.textCritical
  - *@type*
  - *@cause*
  - *@varSeq*
  - *@require*

  - att.typed
    * ~~type~~
    * @subtype

  - att.written
    * *@hand*

*Contained by* *textcrit:* app rdgGrp

*May contain*

*core:* note noteGrp

*textcrit:* lem rdg rdgGrp wit witDetail

*Note* May contain readings and nested reading groups.

Usually, only one <lem> element should appear within a single apparatus entry, whether it appears outside a <rdgGrp> element or within it.

*Example*

```
<app>
  <lem  wit="#El #Ra2">though</lem>
  <rdgGrp  type="orthographic">
    <rdg  wit="#Hg">thogh</rdg>
    <rdg  wit="#La">thouhe</rdg>
```

```
      </rdgGrp>
    </app>
```

*Schematron* <sch:rule context="tei:rdgGrp"> <sch:assert test="count(tei:lem) lt 2">Only one <lem> element may appear within a <rdgGrp></sch:assert> </sch:rule>

*Content model*

```
<content>
 <sequence>
  <elementRef key="lem" minOccurs="0"/>
  <alternate maxOccurs="unbounded"
   minOccurs="0">
   <classRef key="model.rdgLike"/>
   <classRef key="model.noteLike"/>
   <elementRef key="witDetail"/>
   <elementRef key="wit"/>
   <elementRef key="rdgGrp"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element rdgGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.textCritical.attributes,
   att.written.attributes,
   ( lem?, ( model.rdgLike | model.noteLike | witDetail | wit | rdgGrp )* )
}
```

---

**<re>** (related entry) contains a dictionary entry for a lexical item related to the headword, such as a compound phrase or derived form, embedded inside a larger entry. [*10.3.6. Related Entries*]

*Deprecated* will be removed on 2026-03-10

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    - *@datcat*
    - *@valueDatcat*
    - *@targetDatcat*
  - att.lexicographic.normalized
    - *@norm*
    - *@orig*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.entryPart model.entryPart.top
*Contained by*

*core:* cit

*dictionaries:* dictScrap entry entryFree hom re sense

*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q ref reg rs ruby sic soCalled term time title unclear unit

*dictionaries:* def dictScrap entry etym form gramGrp lang lbl oRef pRef re sense usg xr

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May contain character data mixed with any other elements defined in the dictionary tag set.

Identical in sub-elements to an <entry> tag, and used where a dictionary has embedded information inside one entry which could have formed a separate entry. Some authorities distinguish related entries, run-on entries, and various other types of degenerate entries; no such typology is attempted here.

*Example* The following example from *Webster's New Collegiate Dictionary (Springfield, Mass.: G. & C. Merriam Company, 1975)* shows a single related entry for which no definition is given, since its meaning is held to be readily derivable from the root entry:

```
<entry>
  <form>
    <orth>neu·ral</orth>
    <pron>'n(y)ūr-əl</pron>
  </form>
  <gramGrp>
    <pos>adj</pos>
  </gramGrp>
  <sense  n="1">
    <def>of, relating to, or affecting a nerve or the nervous system</def>
  </sense>
  <sense  n="2"> ... </sense>
  <entry  type="relatedEntry">
    <form>
      <orth>neurally</orth>
      <pron  extent="suff">-ə-lē</pron>
```

```
    </form>
    <gramGrp>
      <pos>adv</pos>
    </gramGrp>
  </entry>
</entry>
```

*Example* The following example from *Diccionario de la Universidad de Chicago Inglés-Español y Español-Inglés / The University of Chicago Spanish Dictionary, Fourth Edition, compiled by Carlos Castillo and Otto F. Bond (Chicago: University of Chicago Press, 1987)* shows a number of related entries embedded in the main entry. The original entry resembles the following:

> **abeja** [a·bé·xa]*f.* bee;**abejera** [a·be·xé·ra]*f.* beehive;**abejón** [a·be·xóon]*m.* drone; bumblebee;**abejorro** [a·be·xó·rro]*m.* bumble bee.

One encoding for this entry would be:

```
<entry>
  <form>
    <orth>abeja</orth>
  </form>
  <gramGrp>
    <gen>f. </gen>
  </gramGrp>
  <sense  n="1.">
    <usg  type="domain"> (ento.) </usg>
    <def> bee </def>. </sense>
  <sense  n="2.">
    <def> busy bee, hard worker </def>. </sense>
  <sense  n="3.">
    <usg  orig="A."  type="domain"> (astron.) </usg>, <def> Musca </def> –
</sense>
  <entry  type="relatedEntry">
    <form>
      <orth  orig="a. albanila"> abeja albanila </orth>, </form>
    <sense>
      <def>mason bee</def>;</sense>
  </entry>
  <entry  type="relatedEntry">
    <form>
      <orth  orig="a. carpintera"> abeja carpintera </orth>, </form>
    <sense>
      <def>carpenter bee </def>;</sense>
  </entry>
  <entry  type="relatedEntry">
    <form>
      <orth  xml:id="re-o3"
        orig="a. reina or maestra"> abeja reina </orth>
      <orth  mergedIn="#re-o4"> abeja maestra </orth>
    </form>
    <sense>
      <def> queen bee </def>;</sense>
  </entry>
  <entry  type="relatedEntry">
    <form>
      <orth  xml:id="re-o4"
        orig="a. neutra or obrera"> abeja neutra </orth>
```

```
      <orth mergedIn="#re-o3"> abeja obrera </orth>
    </form>
    <sense>
      <def>worker bee</def>.</sense>
  </entry>
</entry>
```

*Example* In the much larger Simon & Schuster Spanish-English dictionary (*Tana de Gámez, ed., Simon and Schuster's International Dictionary (New York: Simon and Schuster, 1973).*) these derived forms of *abeja* are treated as separate main entries, but there are other embedded phrases shown as <re>s in its main entry for *abeja*:

> abeja, f. 1. (ento.) bee. 2. busy bee, hard worker. 3. (astron.) A., Musca. — a. albanila, mason bee; a. carpintera, carpenter bee; a. reina or maestra, queen bee; a. neutra or obrera, worker bee.

This entry may be encoded thus:

```
<entry>
  <form>
    <orth>abeja</orth>
  </form>
  <gramGrp>
    <gen>f.</gen>
  </gramGrp>
  <sense n="1.">
    <usg type="domain">(ento.)</usg>
    <def>bee</def>. </sense>
  <sense n="2.">
    <def>busy bee, hard worker</def>. </sense>
  <sense n="3.">
    <usg orig="A." type="domain">(astron.)</usg>, <def>Musca</def> — </sense>
  <entry type="relatedEntry">
    <form>
      <orth orig="a. albanila">abeja albanila</orth>, </form>
    <sense>
      <def>mason bee</def>; </sense>
  </entry>
  <entry type="relatedEntry">
    <form>
      <orth orig="a. carpintera">abeja carpintera</orth>, </form>
    <sense>
      <def>carpenter bee</def>; </sense>
  </entry>
  <entry type="relatedEntry">
    <form>
      <orth xml:id="re-o1"
        orig="a. reina or maestra">abeja reina</orth>
      <orth mergedIn="#re-o1">abeja maestra</orth>
    </form>
    <sense>
      <def>queen bee</def>; </sense>
  </entry>
  <entry type="relatedEntry">
    <form>
      <orth xml:id="re-o2"
        orig="a. neutra or obrera">abeja neutra</orth>
```

```
        <orth  mergedIn="#re-o2">abeja obrera</orth>
      </form>
      <sense>
        <def>worker bee</def>. </sense>
    </entry>
  </entry>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <elementRef key="sense"/>
  <classRef key="model.entryPart.top"/>
  <classRef key="model.phrase"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element re
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.typed.attributes,
   (
      text
    | model.gLike    | sense    | model.entryPart.top    | model.phrase    | model.global
}
```

---

**<recordHist>** (recorded history) provides information about the source and revision status of the parent manuscript or object description itself. [*11.9.1. Administrative Information*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Contained by:* *tei:* adminInfo
*May contain*
*core:* p
*header:* change
*linking:* ab
*msdescription:* source
*Example*

```
<recordHist>
  <source>
    <p>Derived from <ref  target="#IMEV">IMEV 123</ref> with additional research
        by P.M.W.Robinson</p>
  </source>
  <change  when="1999-06-23">
    <name>LDB</name> (editor)
     checked examples against DTD version 3.6
  </change>
</recordHist>
```

*Content model*

```
<content>
 <alternate>
```

```
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="source"/>
   <elementRef key="change" minOccurs="0"
    maxOccurs="unbounded"/>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element recordHist
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.pLike+ | ( source, change* ) )
}
```

---

**<recording>** (recording event) provides details of an audio or video recording event used as the source of a spoken text, either directly or from a public broadcast. [*8.2. Documenting the Source of Transcribed Speech 16.3.2. Declarable Elements*]

*Module* spoken — Transcriptions of Speech

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

• att.declarable
– *@default*

• att.duration
– att.duration.iso
* *@dur-iso*

– att.duration.w3c
* *@dur*

• att.typed
– ~~type~~

– @subtype

`@type` the kind of recording.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **audio** audio recording[Default]
**video** audio and video recording

*Contained by:* broadcast recordingStmt
*May contain*
*core:* date media p respStmt time
*linking:* ab
*spoken:* broadcast equipment
*Note* The *dur* attribute is used to indicate the original duration of the recording.
*Example*

```
<recording type="audio" dur="P30M">
  <equipment>
    <p>Recorded on a Sony TR444 walkman by unknown participants; remastered
        to digital tape at <placeName>Borehamwood Studios</placeName> by
    <orgName>Transcription Services Inc</orgName>.</p>
  </equipment>
</recording>
```

*Example*

```
<recording  type="audio"  dur="P10M">
  <equipment>
    <p>Recorded from FM Radio to digital tape</p>
  </equipment>
  <broadcast>
    <bibl>
      <title>Interview on foreign policy</title>
      <author>BBC Radio 5</author>
      <respStmt>
        <resp>interviewer</resp>
        <name>Robin Day</name>
      </respStmt>
      <respStmt>
        <resp>interviewee</resp>
        <name>Margaret Thatcher</name>
      </respStmt>
      <series>
        <title>The World Tonight</title>
      </series>
      <note>First broadcast on
      <date  when="1989-11-27">27 Nov 89</date>
      </note>
    </bibl>
  </broadcast>
</recording>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:recording"/> </sch:pattern>

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.recordingPart"/>
  <classRef key="model.pLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element recording
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   att.typed.attribute.subtype,
   attribute type { "audio" | "video" }?,
   ( model.recordingPart | model.pLike )*
}
```

**<recordingStmt>** (recording statement) describes a set of recordings used as the basis for transcription of a spoken text. [*8.2. Documenting the Source of Transcribed Speech 2.2.7. The Source Description*]

*Module* spoken — Transcriptions of Speech

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.sourceDescPart

*Contained by*

*header:* sourceDesc

*May contain*

*core:* p

*linking:* ab

*spoken:* recording

*Example*

```
<recordingStmt>
  <recording  type="audio"  dur="P30M">
    <respStmt>
      <resp>Location recording by</resp>
      <name>Sound Services Ltd.</name>
    </respStmt>
    <equipment>
      <p>Multiple close microphones mixed down to stereo Digital
            Audio Tape, standard play, 44.1 KHz sampling frequency</p>
    </equipment>
    <date>12 Jan 1987</date>
  </recording>
</recordingStmt>
```

*Example*

```
<recordingStmt>
  <p>Three
     distinct recordings made by hidden microphone in early February
     2001.</p>
</recordingStmt>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <elementRef key="recording" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element recordingStmt
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.pLike+ | recording+ )
}
```

---

**<redo>** indicates one or more cancelled interventions in a document which have subsequently been marked as reaffirmed or repeated. [*12.3.4.4. Confirmation, Cancellation, and Reinstatement of Modifications*]

*Module* transcr — Representation of Primary Sources

*Attributes*
- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

- att.ranging
  - *@atLeast*
  - *@atMost*
  - *@min*
  - *@max*
  - *@confidence*

- att.spanning

- *@spanTo*

- att.transcriptional
  - *@status*
  - *@cause*
  - *@seq*

  - att.editLike
    - *@evidence*
    - *@instant*
  - att.placement
    - *@place*
  - att.written
    - *@hand*

`@target` points to one or more elements representing the interventions which are being reasserted.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Member of* model.linePart model.pPart.transcriptional
*Contained by*
*analysis:* cl pc phr s w
*cmc:* post
*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castItem role roleDesc sound tech view
*figures:* cell
*header:* change distributor edition extent geoDecl handNote licence scriptNote
*linking:* ab seg
*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark
*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname
*spoken:* u writing
*tagdocs:* eg
*textcrit:* lem rdg wit witDetail
*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer
*transcr:* am damage fw line metamark mod restore retrace secl supplied surplus zone
*verse:* rhyme
*May contain* Empty element
*Example*

```
<line>
  <redo  hand="#g_t"  target="#redo-1"
    cause="fix"/>
  <mod  xml:id="redo-1"  rend="strikethrough"
    spanTo="#anchor-1"  hand="#g_bl"/>Ihr hagren, triſten,
krummgezog<mod  rend="strikethrough">nen</mod>ener Nacken
</line>
<line>Wenn ihr nur piepſet iſt die Welt ſchon matt.<anchor  xml:id="anchor-1"/>
</line>
```

This encoding represents the following sequence of events:

- ”Ihr hagren, tristen, krummgezog nenener Nacken/ Wenn ihr nur piepset ist die Welt schon matt.” is written
- the redundant letters ”nen” in ”nenener” are deleted
- the whole passage is deleted by hand `g_bl` using strikethrough
- the deletion is reasserted by another hand (identified here as `g_t`)

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element redo
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.spanning.attributes,
   att.transcriptional.attributes,
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   attribute target { list { + } }?,
   empty
}
```

---

**<ref>** (reference) defines a reference to another location, possibly modified by additional text or comment. [*3.7. Simple Links and Cross-References 17.1. Links*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cReferencing
  - *@cRef*
- att.cmc
  - *@generatedBy*
- att.declaring
  - *@decls*
- att.internetMedia
  - *@mimeType*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.annotationPart.body model.ptrLike
*Contained by*
*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine analytic author bibl biblScope biblStruct cit citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned monogr name note num orig p pubPlace publisher q quote rb ref reg relatedItem resp rs rt said series sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc notatedMusic

*header:* application authority catDesc change classCode correspContext creation distributor edition extent funder geoDecl handNote language licence principal publicationStmt rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab annotation seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education event eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person place placeName region residence roleName settlement sex socecStatus surname

*nets:* eLeaf eTree

*spoken:* transcriptionDesc u writing

*tagdocs:* eg listRef valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink

objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The *target* and *cRef* attributes are mutually exclusive.

*Example*

```
See especially <ref  target="http://www.natcorp.ox.ac.uk/Texts/A02.xml#s2">the
second
sentence</ref>
```

*Example*

```
See also <ref  target="#locution">s.v. <term>locution</term>
</ref>.
```

*Schematron* <sch:rule context="tei:ref"> <sch:report test="@target and @cRef">Only one of the attributes @target and @cRef may be supplied on <sch:name/>.</sch:report> </sch:rule>

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element ref
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cReferencing.attributes,
   att.cmc.attributes,
   att.declaring.attributes,
   att.internetMedia.attributes,
   att.pointing.attributes,
   att.typed.attributes,
   macro.paraContent}
```

**<refsDecl>** (references declaration) specifies how canonical references are constructed for this text. [*2.3.6.3. Milestone Method 2.3. The Encoding Description 2.3.6. The Reference System Declaration*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable
  - *@default*

*Member of* model.encodingDescPart

*Contained by*

*header:* encodingDesc

*May contain*

*core:* p

*header:* cRefPattern citeStructure refState

*linking:* ab

*Example*

```
<refsDecl>
  <cRefPattern  matchPattern="([A-Za-z0-9]+) ([0-9]+):([0-9]+)"
    replacementPattern="#xpath(//body/div[@n='$1']/div[$2]/div3[$3])"/>
</refsDecl>
```

This example is a formal representation for the referencing scheme described informally in the following example.

*Example*

```
<refsDecl>
  <p>References are made up by concatenating the value for the
  <att>n</att> attribute on the highest level <gi>div</gi>
     element, followed by a space, followed by the sequential
     number of the next level <gi>div</gi> followed by a colon
     followed by the sequential number of the next (and lowest)
     level <gi>div</gi>.</p>
</refsDecl>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:refsDecl"/> </sch:pattern>

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <elementRef key="citeStructure"
   minOccurs="1" maxOccurs="unbounded"/>
  <elementRef key="cRefPattern"
   minOccurs="1" maxOccurs="unbounded"/>
  <elementRef key="refState" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element refsDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   ( model.pLike+ | citeStructure+ | cRefPattern+ | refState+ )
}
```

---

**<refState>** (reference state) specifies one component of a canonical reference defined by the milestone method.
[*2.3.6.3. Milestone Method 2.3.6. The Reference System Declaration*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.edition
  - *@ed*
  - *@edRef*
- att.milestoneUnit
  - *@unit*

`@length` specifies the fixed length of the reference component.

*Status* Optional

*Datatype* teidata.count

*Note* When constructing a reference, if the reference component found is of numeric type, the length is made up by inserting leading zeros; if it is not, by inserting trailing blanks. In either case, reference components are truncated if necessary at the right hand side.

When seeking a reference, the length indicates the number of characters which should be compared. Values longer than this will be regarded as matching, if they

start correctly. If no value is provided, the length is unlimited and goes to the next delimiter or to the end of the value.

`@delim` (delimiter) supplies a delimiting string following the reference component.

*Status* Optional

*Datatype* teidata.text

*Contained by:* refsDecl

*May contain* Empty element

*Example*

```
<refState unit="book" delim=":"/>
<refState unit="line" length="4"/>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element refState
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.edition.attributes,
   att.milestoneUnit.attributes,
   attribute length { text }?,
   attribute delim { text }?,
   empty
}
```

---

**<reg>** (regularization) contains a reading which has been regularized or normalized in some sense. [*3.5.2. Regularization and Normalization 13. Critical Apparatus*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*

- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.editLike
  - *@evidence*
  - *@instant*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.choicePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename

genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example* If all that is desired is to call attention to the fact that the copy text has been regularized, <reg> may be used alone:

```
<q>Please <reg>knock</reg> if an <reg>answer</reg> is <reg>required</reg>
</q>
```

*Example* It is also possible to identify the individual responsible for the regularization, and, using the <choice> and <orig> elements, to provide both the original and regularized readings:

```
<q>Please <choice>
    <reg  resp="#LB">knock</reg>
    <orig>cnk</orig>
```

```
      </choice> if an <choice>
        <reg>answer</reg>
        <orig>nsr</orig>
      </choice> is <choice>
        <reg>required</reg>
        <orig>reqd</orig>
      </choice>
    </q>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element reg
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.editLike.attributes,
   att.typed.attributes,
   macro.paraContent}
```

---

**<region>** (region) contains the name of an administrative unit such as a state, province, or county, larger than a settlement, but smaller than a country. [*14.2.3. Place Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*

- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
    - *@key*
    - *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.placeNamePart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage unitDef

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions altIdentifier catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msIdentifier musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectIdentifier objectName occupation offset org orgName persName persPronouns place placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<placeName>
  <region  type="state"  n="IL">Illinois</region>
</placeName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element region
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

**<relatedItem>** contains or references some other bibliographic item which is related to the present one in some specified manner, for example as a constituent or alternative version of it. [*3.12.2.7. Related Items*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - *@type*
  - *@subtype*

`@target` points to the related bibliographic element by means of an absolute or relative URI reference.
*Status* Optional
*Datatype* teidata.pointer

*Member of* model.biblPart

*Contained by*
*core:* bibl biblStruct
*header:* notesStmt
*May contain*
*core:* bibl biblStruct listBibl ptr ref
*header:* biblFull
*msdescription:* msDesc
*tagdocs:* listRef
*Note* If the *target* attribute is used to reference the related bibliographic item, the element must be empty.
*Example*

```
<biblStruct>
  <monogr>
    <author>Shirley, James</author>
    <title  type="main">The gentlemen of Venice</title>
    <imprint>
      <pubPlace>New York</pubPlace>
      <publisher>Readex Microprint</publisher>
      <date>1953</date>
    </imprint>
    <extent>1 microprint card, 23 x 15 cm.</extent>
  </monogr>
  <series>
    <title>Three centuries of drama: English, 1642–1700</title>
  </series>
  <relatedItem  type="otherForm">
    <biblStruct>
      <monogr>
        <author>Shirley, James</author>
        <title  type="main">The gentlemen of Venice</title>
        <title  type="sub">a tragi-comedie presented at the private house in
Salisbury
               Court by Her Majesties servants</title>
        <imprint>
          <pubPlace>London</pubPlace>
          <publisher>H. Moseley</publisher>
          <date>1655</date>
        </imprint>
        <extent>78 p.</extent>
      </monogr>
    </biblStruct>
  </relatedItem>
</biblStruct>
```

*Schematron* <sch:rule context="tei:relatedItem"> <sch:report test="@target and count( child::* ) > 0">If the @target attribute on <sch:name/> is used, the relatedItem element must be empty</sch:report> <sch:assert test="@target or child::*">A relatedItem element should have either a @target attribute or a child element to indicate the related bibliographic item</sch:assert> </sch:rule>
*Content model*

```
<content>
 <alternate minOccurs="0">
  <classRef key="model.biblLike"/>
  <classRef key="model.ptrLike"/>
 </alternate>
```

```
</content>
```

*Schema Declaration*

```
element relatedItem
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   attribute target { text }?,
   ( model.biblLike | model.ptrLike )?
}
```

---

**<relation>** (relationship) describes any kind of relationship or linkage amongst a specified group of places, events, persons, objects or other items. [*14.3.2.3. Personal Relationships*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

`@name` supplies a name for the kind of relationship of which this is an instance.

*Status* Optional

*Datatype* teidata.enumerated

`@active` identifies the 'active' participants in a non-mutual relationship, or all the participants in a mutual one.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

`@mutual` supplies a list of participants amongst all of whom the relationship holds equally.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

`@passive` identifies the 'passive' participants in a non-mutual relationship.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Contained by* *core:* listBibl

*namesdates:* event listEvent listNym listObject listOrg listPerson listPlace listRelation

*May contain*

*core:* desc

*Note* Only one of the attributes *active* and *mutual* may be supplied; the attribute *passive* may be supplied only if the attribute *active* is supplied. Not all of these constraints can be enforced in all schema languages.

*Example*

```
<relation type="social" name="supervisor"
  active="#p1" passive="#p2 #p3 #p4"/>
```

This indicates that the person with identifier p1 is supervisor of persons p2, p3, and p4.

*Example*

```
<relation type="personal" name="friends"
  mutual="#p2 #p3 #p4"/>
```

This indicates that p2, p3, and p4 are all friends.

*Example*

```
<relation type="CRM"
  name="P89_falls_within"
  active="http://id.clarosnet.org/places/metamorphoses/place/italy-orvieto"
  passive="http://id.clarosnet.org/places/metamorphoses/country/IT"/>
```

This indicates that there is a relation, defined by CIDOC CRM, between two resources identified by URLs.

*Example*

```
<relation resp="http://viaf.org/viaf/44335536/"
  ref="http://purl.org/saws/ontology#isVariantOf"
  active="http://www.ancientwisdoms.ac.uk/cts/urn:cts:greekLit:tlg3017.Syno298.sawsGrc01:dive
  passive="http://data.perseus.org/citations/urn:cts:greekLit:tlg0031.tlg002.perseus-grc1:9.3
```

This example records a relationship, defined by the SAWS ontology, between a passage of text identified by a CTS URN, and a variant passage of text in the Perseus Digital Library, and assigns the identification of the relationship to a particular editor (all using resolvable URIs).

*Schematron* <sch:rule context="tei:relation"> <sch:assert test="@ref or @key or @name">One of the attributes @name, @ref or @key must be supplied</sch:assert> </sch:rule>

*Schematron* <sch:rule context="tei:relation"> <sch:report test="@active and @mutual">Only one of the attributes @active and @mutual may be supplied</sch:report> </sch:rule>

*Schematron* <sch:rule context="tei:relation"> <sch:report test="@passive and not(@active)">the attribute @passive may be supplied only if the attribute @active is supplied</sch:report> </sch:rule>

*Content model*

```
<content>
 <elementRef key="desc" minOccurs="0"/>
</content>
```

*Schema Declaration*

```
element relation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   attribute name { text }?,
   ( attribute active { list { + } }? | attribute mutual { list { + } }? ),
   attribute passive { list { + } }?,
   desc?
}
```

---

## <remarks>

**<remarks>** (remarks) contains any commentary or discussion about the usage of an element, attribute, class, or entity not otherwise documented within the containing element. [*23.5. Element Specifications 23.5.3. Attribute List Specification 23.6. Class Specifications 23.7. Macro Specifications*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change

* *@change*

– att.global.facs
* *@facs*

– att.global.linking
* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

• att.combinable
– *@mode*

– att.deprecated
* *@validUntil*

• att.translatable
– *@versionDate*

`@ident` supplies the identifier by which the remarks may be referenced.
*Status* Optional
*Datatype* teidata.name

*Contained by* tagdocs: attDef classSpec dataSpec elementSpec macroSpec moduleSpec valItem
*May contain*
*core:* p
*linking:* ab
*Example*

```
<remarks>
  <p>This element is probably redundant.</p>
</remarks>
```

*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="1"
```

```
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element remarks
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   att.translatable.attributes,
   attribute ident { text }?,
   model.pLike+
}
```

---

**<rendition>** (rendition) supplies information about the rendition or appearance of one or more elements in the source text. [*2.3.4. The Tagging Declaration*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

  * *@select*
  - att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
  - att.global.responsibility
  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.styleDef
  - *@scheme*
  - *@schemeVersion*

`@scope` where CSS is used, provides a way of defining 'pseudo-elements', that is, styling rules applicable to specific sub-portions of an element.
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **first-line** styling applies to the first line of the target element
**first-letter** styling applies to the first letter of the target element
**before** styling should be applied immediately before the content of the target element
**after** styling should be applied immediately after the content of the target element

`@selector` contains a selector or series of selectors specifying the elements to which the contained style description applies, expressed in the language specified in the *scheme* attribute.
*Status* Optional
*Datatype* teidata.text

```
<rendition scheme="css"
  selector="text, front, back, body, div, p, ab">
display: block;
</rendition>
<rendition scheme="css"
  selector="*[rend*=italic]"> font-style: italic;
</rendition>
```

*Note* Since the default value of the *scheme* attribute is assumed to be CSS, the default expectation for this attribute, in the absence of *scheme*, is that CSS selector syntax will be used.
While *rendition* is used to point from an element in the transcribed source to a <rendition> element in the header which describes how it appears, the *selector* attribute allows the encoder to point in the other direction: from a <rendition> in the header to a collection of elements which all share the same renditional features. In both cases, the intention is to record the appearance of the source text, not to prescribe any particular output rendering.

*Contained by:* tagsDecl
*May contain*
*core:* abbr address bibl biblStruct choice cit date desc distinct email emph expan foreign gloss hi label list

listBibl measure measureGrp mentioned name num ptr q quote ref rs said soCalled stage term time title unit

*dictionaries:* lang

*drama:* camera caption castList move sound tech view

*figures:* table

*header:* biblFull idno

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specGrp specGrpRef tag val

*textcrit:* listApp listWit

*textstructure:* floatingText

*transcr:* am ex subst

character data

*Example*

```
<tagsDecl>
  <rendition  xml:id="r-center"  scheme="css">text-align: center;</rendition>
  <rendition  xml:id="r-small"  scheme="css">font-size: small;</rendition>
  <rendition  xml:id="r-large"  scheme="css">font-size: large;</rendition>
  <rendition  xml:id="initcaps"
    scope="first-letter"  scheme="css">font-size: xx-large</rendition>
</tagsDecl>
```

*Content model*

```
<content>
 <macroRef key="macro.limitedContent"/>
</content>
```

*Schema Declaration*

```
element rendition
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.styleDef.attributes,
   attribute scope { text }?,
   attribute selector { text }?,
   macro.limitedContent}
```

**<repository>** (repository) contains the name of a repository within which manuscripts or other objects are stored, possibly forming part of an institution. [*11.4. The Manuscript Identifier*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
    - *@key*
    - *@ref*

*Contained by:* *msdescription:* altIdentifier msIdentifier
*namesdates:* objectIdentifier

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin

character data

*Example*

```
<msIdentifier>
  <settlement>Oxford</settlement>
  <institution>University of Oxford</institution>
  <repository>Bodleian Library</repository>
  <idno>MS. Bodley 406</idno>
</msIdentifier>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element repository
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   macro.phraseSeq.limited}
```

**<residence>** (residence) describes a person's present or past places of residence. [*16.2.2. The Participant Description*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*
    - *@to-custom*
    - *@datingPoint*

  * *@datingMethod*
  - att.datable.iso
  * *@when-iso*
  * *@notBefore-iso*
  * *@notAfter-iso*
  * *@from-iso*
  * *@to-iso*
  - att.datable.w3c
  * *@when*
  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*

  - att.canonical
  * *@key*
  * *@ref*
- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **primary**
**secondary**
**temporary**
**permanent**

*Member of* model.persStateLike
*Contained by*
*namesdates:* person personGrp persona
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<residence>Childhood in East Africa and long term resident of Glasgow,
Scotland.</residence>
```

*Example*

```
<residence  notAfter="1997">Mbeni estate, Dzukumura region, Matabele
land</residence>
<residence  notBefore="1903"  notAfter="1996">
  <placeName>
    <settlement>Glasgow</settlement>
    <region>Scotland</region>
  </placeName>
</residence>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element residence
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
```

```
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.phraseSeq}
```

---

**<resp>** (responsibility) contains a phrase describing the nature of a person's intellectual responsibility, or an organization's role in the production or distribution of a work. [*3.12.2.2. Titles, Authors, and Editors 2.2.1. The Title Statement 2.2.2. The Edition Statement 2.2.5. The Series Statement*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

- att.canonical
  - *@key*
  - *@ref*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*
    - *@to-custom*
    - *@datingPoint*
    - *@datingMethod*
  - att.datable.iso
    - *@when-iso*
    - *@notBefore-iso*
    - *@notAfter-iso*
    - *@from-iso*
    - *@to-iso*
  - att.datable.w3c
    - *@when*
    - *@notBefore*
    - *@notAfter*
    - *@from*
    - *@to*

*Contained by* *edrb* respStmt

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data

*Note* The attribute *ref*, inherited from the class `att.canonical` may be used to indicate the kind of responsibility in a normalized form by referring directly to a standardized list of responsibility types, such as that maintained by a naming authority, for example the list maintained at `http://www.loc.gov/marc/relators/relacode.html` for bibliographic usage.

*Example*

```
<respStmt>
  <resp  ref="http://id.loc.gov/vocabulary/relators/com.html">compiler</resp>
  <name>Edward Child</name>
</respStmt>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element resp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   macro.phraseSeq.limited}
```

---

**<respons>** (responsibility) identifies the individual(s) responsible for some aspect of the content or markup of particular element(s). [*22.3. Attribution of Responsibility*]

*Module* certainty — Certainty, Precision, and Responsibility

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*

- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.scoping
  - *@target*
  - *@match*

`@locus` indicates the specific aspect of the encoding (markup or content) for which responsibility is being assigned.
*Status* Required
*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace
*Legal values are:* **name** responsibility is being assigned concerning the name of the element or attribute used.
**start** responsibility is being assigned concerning the start of the element concerned.
**end** responsibility is being assigned concerning the end of the element concerned.
**location** responsibility is being assigned concerning the location of the element concerned.
**value** responsibility is being assigned concerning the content (for an element) or the value (for an attribute)

*Member of* model.certLike model.global.meta
*Contained by*
*analysis:* cl interp m phr s span w
*certainty:* certainty precision respons
*cmc:* post
*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gap gloss head headItem headLabel hi imprint item l label lg list measure mentioned name

note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab join seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc space substJoin supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*certainty:* certainty precision respons

*core:* desc

*Note* The <respons> element is designed for cases in which fine-grained information about specific aspects of the markup of a text is desirable for whatever reason. Global responsibility for certain aspects of markup is usually more simply indicated in the TEI header, using the <respStmt> element within the title statement, edition statement, or change log.

*Example*

```
<respons target="#p1" locus="name location"
  resp="#encoder1"/>
<respons target="#p2" match="@rend"
  locus="value" resp="#encoder2"/>
<list type="encoders">
  <item xml:id="encoder1"/>
  <item xml:id="encoder2"/>
</list>
```

*Example* In this (partially fictional) example the entire document was transcribed and encoded by a single encoder, except for one passage which was transcribed and encoded by the proofreader.

```
<!-- in the <teiHeader>: --><respStmt  xml:id="enc01">
  <name>C. Colin Backslash</name>
  <resp>transcription</resp>
  <resp>encoding</resp>
</respStmt>
<respStmt  xml:id="prf01">
  <name>Erin Spelling</name>
  <resp>proofreading</resp>
</respStmt>
<!-- in the <text>: -->
<p>θεριν‍ὴ τροπή
<lb/>
  <foreign  xml:lang="hbo"  xml:id="mp0a8">
  <supplied  reason="undefined">□□</supplied>
    <unclear/>□<unclear>
     □ □□□□
  </foreign>
</p>
<!-- elsewhere: -->
<respons  target="#mp0a8"  locus="name value"
  resp="#prf01"/>
```

*Example* In this (partially fictional) example an initial encoder (‘across.dta’) encoded most of the document; a later encoder (‘rcapolung.ewo’) encoded a particular passage; and a third encoder (‘sbauman.emt’) fixed some of that encoding.

```
<!-- in the <teiHeader>: --><respStmt  xml:id="enc03">
  <persName  ref="../contextual/persons.xml#across.dta"/>
  <resp>encoding</resp>
</respStmt>
<!-- in the <text>: -->
<spGrp  rend="braced(atonce1)"
  xml:id="sgrp05">
  <sp  who="#mo">
    <speaker  rend="align(left)slant(italic)">
      <persName>Mor</persName>.</speaker>
    <p  rend="break(no)">So, so, so!</p>
  </sp>
  <sp  who="#hg">
    <speaker  rend="align(left)slant(italic)">
      <persName>Mr. H</persName>.</speaker>
    <p  rend="break(no)">What, without my Leave!</p>
  </sp>
  <sp  who="#la">
    <speaker  rend="align(left)slant(italic)">
      <persName>Lady D</persName>.</speaker>
    <p  rend="break(no)">Amazing!</p>
  </sp>
</spGrp>
<stage  rend="align(right)slant(italic)"
  type="delivery"  xml:id="atonce1">All together.</stage>
<!-- anywhere: -->
<respons  target="sgrp05"  locus="name"
  resp="../contextual/persons.xml#rcapolung.ewo">
  <desc>Ashley did not know what to do with this; I have decided it
     best fits as a braced <gi>spGrp</gi>
```

```
  </desc>
</respons>
<respons  target="sgrp05"  match=".//@rend"
  locus="value"
  resp="../contextual/persons.xml#sbauman.emt">
  <desc>fixed <att>rend</att> attributes</desc>
</respons>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.descLike"/>
  <classRef key="model.certLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element respons
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.scoping.attributes,
   attribute locus
   {
      list { ( "name" | "start" | "end" | "location" | "value" )+ }
   },
   ( model.descLike | model.certLike )*
}
```

---

**<respStmt>** (statement of responsibility) supplies a statement of responsibility for the intellectual content of a text, edition, recording, or series, where the specialized elements for authors, editors, etc. do not suffice or do not apply. May also be used to encode information about individuals or organizations which have played a role in the production or distribution of a bibliographic work. [*3.12.2.2. Titles, Authors, and Editors 2.2.1. The Title Statement 2.2.2. The Edition Statement 2.2.5. The Series Statement*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.canonical
  - *@key*
  - *@ref*

*Member of* model.recordingPart model.respLike
*Contained by*
*core:* analytic bibl imprint monogr series
*header:* editionStmt seriesStmt titleStmt
*linking:* annotation
*msdescription:* msItem msItemStruct
*spoken:* recording
*May contain*
*core:* name note resp
*namesdates:* orgName persName
*Example*

```
<respStmt>
  <resp>transcribed from original ms</resp>
  <persName>Claus Huitfeldt</persName>
</respStmt>
```

*Example*

```
<respStmt>
  <resp>converted to XML encoding</resp>
  <name>Alan Morrison</name>
</respStmt>
```

*Content model*

```
<content>
 <sequence>
  <alternate>
   <sequence>
    <elementRef key="resp" minOccurs="1"
     maxOccurs="unbounded"/>
    <classRef key="model.nameLike.agent"
     minOccurs="1" maxOccurs="unbounded"/>
   </sequence>
   <sequence>
    <classRef key="model.nameLike.agent"
     minOccurs="1" maxOccurs="unbounded"/>
    <elementRef key="resp" minOccurs="1"
     maxOccurs="unbounded"/>
   </sequence>
  </alternate>
  <elementRef key="note" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element respStmt
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   (
      ( ( resp+, model.nameLike.agent+ ) | ( model.nameLike.agent+, resp+ ) ),
      note*
   )
}
```

---

## <restore>

<restore> (restore) indicates restoration of text to an earlier state by cancellation of an editorial or authorial marking or instruction. [*12.3.1.6. Cancellation of Deletions and Other Markings*]

*Module* transcr — Representation of Primary Sources

*Attributes*
- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

- att.ranging
  - *@atLeast*
  - *@atMost*
  - *@min*
  - *@max*
  - *@confidence*
- att.transcriptional
  - *@status*

  - *@cause*
  - *@seq*

  - att.editLike
    * *@evidence*
    * *@instant*
  - att.placement
    * *@place*
  - att.written
    * *@hand*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.linePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media

mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* On this element, the *type* attribute categorizes the way that the cancelled intervention has been indicated in some way, for example by means of a marginal note, over-inking, additional markup, etc.

*Example*

```
For I hate this
<restore  hand="#dhl"
  type="marginalStetNote">
  <del>my</del>
</restore> body
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element restore
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
```

```
   att.dimensions.attributes,
   att.ranging.attributes,
   att.transcriptional.attributes,
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   att.typed.attributes,
   macro.paraContent}
```

---

**<retrace>** contains a sequence of writing which has been retraced, for example by over-inking, to clarify or fix it. [*12.3.4.3. Fixation and Clarification*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.dimensions

- *@unit*
- *@quantity*
- *@extent*
- *@precision*
- *@scope*

- att.ranging
  - *@atLeast*
  - *@atMost*
  - *@min*
  - *@max*
  - *@confidence*
- att.spanning
  - *@spanTo*
- att.transcriptional
  - *@status*
  - *@cause*
  - *@seq*

- att.editLike
  - *@evidence*
  - *@instant*
- att.placement
  - *@place*
- att.written
  - *@hand*

*Member of* model.linePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename

genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Multiple retraces are indicated by nesting one <retrace> within another. In principle, a retrace differs from a substitution in that second and subsequent rewrites do not materially alter the content of an element. Where minor changes have been made during the retracing action however these may be marked up using <del>, <add>, etc. with an appropriate value for the *change* attribute.

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element retrace
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.spanning.attributes,
   att.transcriptional.attributes,
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   macro.paraContent}
```

---

**<revisionDesc>** (revision description) summarizes the revision history for a file. [*2.6. The Revision Description 2.1.1. The TEI Header and Its Components*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

- att.docStatus
  - *@status*

*Contained by:* *header:* teiHeader
*linking:* annotation
*May contain*
*core:* list
*header:* change listChange
*Note* If present on this element, the *status* attribute should indicate the current status of the document. The same attribute may appear on any <change> to record the status at the time of that change. Conventionally <change> elements should be given in reverse date order, with the most recent change at the start of the list.
*Example*

```
<revisionDesc status="embargoed">
  <change when="1991-11-11" who="#LB"> deleted chapter 10 </change>
</revisionDesc>
```

*Content model*

```
<content>
 <alternate>
  <elementRef key="list" minOccurs="1"
   maxOccurs="unbounded"/>
  <elementRef key="listChange"
   minOccurs="1" maxOccurs="unbounded"/>
  <elementRef key="change" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element revisionDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.docStatus.attributes,
```

```
   ( list+ | listChange+ | change+ )
}
```

---

**<rhyme>** marks the rhyming part of a metrical line. [*6.5. Rhyme*]

*Module* verse — Verse

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - *@type*
  - *@subtype*

`@label` provides a label (usually a single letter) to identify which part of a rhyme scheme this rhyming string instantiates.

*Status* Recommended

*Datatype* teidata.word

*Note* Within a particular scope, all <rhyme> elements with the same value for their *label* attribute are assumed to rhyme with each other. The scope is defined by the nearest ancestor element for which the *rhyme* attribute has been supplied.

*Member of* model.lPart

*Contained by*

*analysis:* cl phr s w

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat

geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<lg  rhyme="abababcc">
  <l>'Tis pity learned virgins ever <rhyme  label="a">wed</rhyme>
  </l>
  <l>With persons of no sort of edu<rhyme  label="b">cation</rhyme>,</l>
  <l>Or gentlemen, who, though well born and
<rhyme  label="a">bred</rhyme>,</l>
  <l>Grow tired of scientific conver<rhyme  label="b">sation</rhyme>:</l>
  <l>I don't choose to say much on this <rhyme  label="a">head</rhyme>,</l>
  <l>I'm a plain man, and in a single <rhyme  label="b">station</rhyme>,</l>
  <l>But — Oh! ye lords of ladies inte<rhyme  label="c">llectual</rhyme>,</l>
  <l>Inform us truly, have they not hen-<rhyme  label="c">peck'd you
all</rhyme>?</l>
</lg>
```

*Example*

```
<lg>
  <l>Tyger! Tyger! burning <rhyme  label="a">bright</rhyme>
  </l>
  <l>In the forests of the <rhyme  label="a">night</rhyme>,</l>
  <l>What immortal hand or <rhyme  label="b">eye</rhyme>
  </l>
  <l>Could frame thy fearful
<rhyme  label="b"  type="eye-rhyme">symmetry</rhyme>?</l>
</lg>
```

*Example*

```
<lg>
  <l>"Hark! Lakshman! Hark, again that <rhyme  label="a">cry</rhyme>!</l>
  <l>It is, — it is my husband's <rhyme  label="b">voice</rhyme>!</l>
  <l>hasten, to his succour <rhyme  label="a">fly</rhyme>,</l>
  <l>No more hast thou, dear friend, a <rhyme  label="b">choice</rhyme>.</l>
  <l>He calls on thee, perhaps his <rhyme  label="c">foes</rhyme>
  </l>
  <l>Environ him on all sides <rhyme  label="d">round</rhyme>,</l>
  <l>That wail, — it means death's final <rhyme  label="c">throes</rhyme>!</l>
  <l>Why standest thou, as magic-<rhyme  label="d">bound</rhyme>?</l>
</lg>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element rhyme
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   attribute label { text }?,
   macro.paraContent}
```

---

**<role>** (role) contains the name of a dramatic role, as given in a cast list. [*7.1.4. Cast Lists*]

*Module* drama — Performance Texts

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition

* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@gender` specifies the gender of the role.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.gender separated by whitespace
*Note* Values for this attribute may be locally defined by a project, or they may refer to an external standard.

*Member of* model.castItemPart
*Contained by*
*drama:* castItem
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit
*dictionaries:* lang oRef pRef
*figures:* figure formula notatedMusic
*gaiji:* g
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef specDesc specList tag val
*textcrit:* app witDetail
*textstructure:* floatingText
*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo
*verse:* caesura rhyme
character data
*Note* It is important to assign a meaningful ID attribute to the <role> element, since this ID is referred to by *who* attributes on many other elements.
*Example*

```
<role xml:id="jt">Joan Trash</role>
<roleDesc>A Ginger-bread-woman</roleDesc>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element role
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute gender { list { + } }?,
   macro.phraseSeq}
```

---

## <roleDesc>

(role description) describes a character's role in a drama. [*7.1.4. Cast Lists*]

*Module* drama — Performance Texts

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Member of* model.castItemPart

*Contained by*

*drama:* castGroup castItem

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<roleDesc>gentlemen of leisure</roleDesc>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element roleDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

---

## <roleName>

(role name) contains a name component which indicates that the referent has a particular role or position in society, such as an official title or rank. [*14.2.1. Personal Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*

- att.global.source
  - *@source*
- att.cmc
  - *@generatedBy*
- att.personal
  - *@full*
  - *@sort*
- att.naming
  - *@role*
  - *@nymRef*
- att.canonical
  - *@key*
  - *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.persNamePart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* A <roleName> may be distinguished from an <addName> by virtue of the fact that, like a title, it typically exists independently of its holder.

*Example*

```
<persName>
  <forename>William</forename>
  <surname>Poulteny</surname>
  <roleName>Earl of Bath</roleName>
</persName>
```

*Example*

```
<p>The <roleName  role="solicitor_general">S.G.</roleName> is the only national
public official,
including the Supreme Court justices, required by statute to be “learned in the
law.”</p>
```

*Example*

```
<p>
  <persName  ref="#NJF">
    <roleName  role="solicitor_general">Solicitor General</roleName> Noel J.
```

```
Francisco</persName>,
representing the administration, asserted in rebuttal that there was nothing to
disavow (...)
<persName  ref="#NJF">Francisco</persName> had violated the scrupulous standard
of candor about the facts and
the law that <roleName  role="solicitor_general">S.G.s</roleName>, in
Republican and Democratic administrations
alike, have repeatedly said they must honor.
</p>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element roleName
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.personal.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<root>** (root node) represents the root node of a tree. [*20.2. Trees*]

*Module* nets — Graphs, Networks, and Trees

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs

* *@facs*

– att.global.linking
* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@value` identifies the root node of the network by pointing to a feature structure or other analytic element.
*Status* Optional
*Datatype* teidata.pointer

`@children` identifies the elements which are the children of the root node.
*Status* Required
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace
*Note* If the root has no children (i.e., the tree is 'trivial'), then the *children* attribute should be omitted.

`@ord` (ordered) indicates whether or not the root is ordered.
*Status* Optional
*Datatype* teidata.xTruthValue
*Note* The value true indicates that the children of the root are ordered, whereas false indicates the are unordered.
Use if and only if *ord* is specified as partial on the <tree> element and the root has more than one child.

`@outDegree` gives the out degree of the root, the number of its children.
*Status* Optional
*Datatype* teidata.count
*Note* The in degree of the root is always 0.

*Contained by* tree
*May contain*
*core:* label
*Example*

```
<root  xml:id="vp1"  children="#vb1 #pn1"
  outDegree="2">
  <label>VP</label>
</root>
<leaf  xml:id="vb1"/>
<leaf  xml:id="pn1"/>
```

*Content model*

```
<content>
 <elementRef key="label" minOccurs="0"/>
</content>
```

*Schema Declaration*

```
element root
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute value { text }?,
   attribute children { list { + } },
   attribute ord { text }?,
   attribute outDegree { text }?,
   label?
}
```

---

**<row>** (row) contains one row of a table. [*15.1.1. TEI Tables*]

*Module* figures — Tables, Formulæ, Graphics, and Notated Music

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*

  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

- att.tableDecoration
  - *@role*
  - *@rows*
  - *@cols*

*Contained by:* table
*May contain figures:* cell
*Example*

```
<row  role="data">
  <cell  role="label">Classics</cell>
  <cell>Idle listless and unimproving</cell>
</row>
```

*Content model*

```
<content>
 <elementRef key="cell" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element row
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.tableDecoration.attributes,
   cell+
}
```

---

**<rs>** (referencing string) contains a general purpose name or referring string. [*14.2.1. Personal Names 3.6.1. Referring Strings*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.naming
  - *@role*

  - *@nymRef*

  - att.canonical
    - *@key*
    - *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.nameLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msName musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef specDesc specList tag val
*textcrit:* app witDetail
*textstructure:* floatingText
*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo
*verse:* caesura rhyme
character data

*Example*

```
<q>My dear <rs  type="person">Mr. Bennet</rs>, </q> said <rs  type="person">his
lady</rs>
to him one day,
<q>have you heard that <rs  type="place">Netherfield Park</rs> is let at
last?</q>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element rs
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<rt>** (ruby text) contains a ruby text, an annotation closely associated with a passage of the main text. [*3.4.2. Ruby Annotations*]

*Module* core — Elements Available in All TEI Documents
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  * *@ana*
- att.global.change
  * *@change*
- att.global.facs
  * *@facs*
- att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

• att.transcriptional
- *@status*
- *@cause*
- *@seq*

- att.editLike
  * *@evidence*
  * *@instant*
- att.placement
  * *@place*
- att.written

  * *@hand*

- att.typed
  - *@type*
  - *@subtype*

`@target` supplies a pointer to the base being glossed by this ruby text.
*Status* Optional
*Datatype* teidata.pointer
*Schematron* Enforce that *either target* or both *from* and *to* (or none) are used, but not *target* in combination with either *from* or *to*. <sch:rule context="tei:rt/@target"> <sch:report test="../@from | ../@to">When target= is present, neither from= nor to= should be.</sch:report> </sch:rule>
*Note* Should point to a single <rb> or an element that is inside an <rb>. To refer to multiple elements or text nodes at once use *from* and *to*.

`@from` points to the starting point of the span of text being glossed by this ruby text.
*Status* Optional
*Datatype* teidata.pointer
*Schematron* Enforce the presence of *to* iff there is a *from*. <sch:rule context="tei:rt/@from"> <sch:assert test="../@to">When from= is present, the to= attribute of <sch:name/> is required.</sch:assert> </sch:rule>

`@to` points to the ending point of the span of text being glossed.
*Status* Optional
*Datatype* teidata.pointer
*Schematron* Enforce the presence of *from* iff there is a *to*. <sch:rule context="tei:rt/@to"> <sch:assert test="../@from">When to= is present, the from= attribute of <sch:name/> is required.</sch:assert> </sch:rule>

*Contained by:* ruby
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit
*dictionaries:* lang oRef pRef
*figures:* figure formula notatedMusic
*gaiji:* g
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef specDesc specList tag val
*textcrit:* app witDetail
*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Where the *place* attribute is not provided on the <rt> element, the default assumption is that the ruby gloss is above where the text is horizontal, and to the right of the text where it is vertical.

*Example* The word □□□ *daitōryō* (president) is glossed character by character in hiragana to provide a pronunciation guide.

```
<p  style="writing-mode: vertical-rl"
  xml:lang="ja">
<!--…-->
  <ruby>
    <rb>大</rb>
    <rt  place="right">だい</rt>
  </ruby>
  <ruby>
    <rb>統</rb>
    <rt  place="right">とう</rt>
  </ruby>
  <ruby>
    <rb>領</rb>
    <rt  place="right">りょう</rt>
  </ruby>
<!--…-->
</p>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element rt
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.transcriptional.attributes,
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   att.typed.attributes,
   attribute target { text }?,
   attribute from { text }?,
   attribute to { text }?,
   macro.phraseSeq}
```

**<rubric>** (rubric) contains the text of any *rubric* or heading attached to a particular manuscript item, that is, a string of words through which a manuscript or other object signals the beginning of a text division, often with an assertion as to its author and title, which is in some way set off from the text itself, typically in red ink, or by use of different size or type of script, or some other such visual device. [*11.6.1. The msItem and msItemStruct Elements*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.msExcerpt
  - *@defective*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.msQuoteLike

*Contained by*

*msdescription:* msItem msItemStruct

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<rubric>Nu koma Skyckiu Rym<ex>ur</ex>.</rubric>
<rubric>Incipit liber de consciencia humana a beatissimo Bernardo
editus.</rubric>
<rubric>
  <locus>16. f. 28v in margin: </locus>Dicta Cassiodori
</rubric>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element rubric
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.msExcerpt.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<ruby>** (ruby container) contains a passage of base text along with its associated ruby gloss(es). [*3.4.2. Ruby Annotations*]

*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.typed
  - *@type*

– *@subtype*

*Member of* model.phrase

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* rb rt

*Example* The word 入学試験 *nyūgakushiken* (university entrance exam) is glossed with a hiragana phonation guide.

```
<p  xml:lang="ja">
<!--…-->
  <ruby>
    <rb>入学試験</rb>
    <rt  place="above">にゅうがくしけん</rt>
  </ruby>
<!--…-->
</p>
```

*Example* This fictional example shows the initialism *TEI* glossed letter-by-letter with an IPA transcription.

```
<ruby>
  <rb>T</rb>
  <rt>ti:</rt>
</ruby>
<ruby>
  <rb>E</rb>
  <rt>i:</rt>
</ruby>
```

```
<ruby>
  <rb>I</rb>
  <rt>aɪ</rt>
</ruby>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="rb" minOccurs="1"
   maxOccurs="1"/>
  <elementRef key="rt" minOccurs="1"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element ruby
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.typed.attributes,
   ( rb, rt+ )
}
```

---

**<S>** (s-unit) contains a sentence-like division of a text. [*18.1. Linguistic Segment Categories 8.4.1. Segmentation*]

*Module* analysis — Simple Analytic Mechanisms

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.notated
  - *@notation*
- att.segLike
  - *@function*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.fragmentable
    * *@part*
  - att.metrical
    * *@met*
    * *@real*
    * *@rhyme*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.segLike

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p

pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The <s> element may be used to mark orthographic sentences, or any other segmentation of a text,

provided that the segmentation is end-to-end, complete, and non-nesting. For segmentation which is partial or recursive, the <seg> should be used instead.

The *type* attribute may be used to indicate the type of segmentation intended, according to any convenient typology.

*Example*

```
<head>
  <s>A short affair</s>
</head>
<s>When are you leaving?</s>
<s>Tomorrow.</s>
```

*Schematron* <sch:rule context="tei:s"> <sch:report test="tei:s">You may not nest one s element within another: use seg instead</sch:report> </sch:rule>

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element s
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.notated.attributes,
   att.segLike.attributes,
   att.datcat.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<said>** (speech or thought) indicates passages thought or spoken aloud, whether explicitly indicated in the source or not, whether directly or indirectly reported, whether by real people or fictional characters. [*3.3.3. Quotation*]

*Module* core — Elements Available in All TEI Documents

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.ascribed.directed
  - *@toWhom*
  - att.ascribed
    * *@who*
- att.cmc
  - *@generatedBy*

`@aloud` may be used to indicate whether the quoted matter is regarded as having been vocalized or signed.

*Status* Optional

*Datatype* teidata.xTruthValue

```
<p> Celia thought privately, <said  aloud="false">Dorothea quite
despises Sir James Chettam;
    I believe she would not accept him.</said> Celia felt that
this was a pity.

<!-- ... -->
</p>
```

*Note* The value true indicates the encoded passage was expressed outwardly (whether spoken, signed, sung, screamed, chanted, etc.); the value false indicates that the encoded passage was thought, but not outwardly expressed.

`@direct` may be used to indicate whether the quoted matter is regarded as direct or indirect speech.

*Status* Optional

*Datatype* teidata.xTruthValue

*Default* true

```
<!-- in the header --><editorialDecl>
  <quotation marks="none"/>
</editorialDecl>
<!-- ... -->
<p>Tantripp had brought a card, and said that
<said direct="false">there was a gentleman waiting in the
lobby</said>.
The courier had told him that <said direct="false">only Mrs.
Casaubon was at home</said>,
but he said <said direct="false">he was a relation of Mr.
Casaubon's: would she see him?</said>
</p>
```

*Note* The value true indicates the speech or thought is represented directly; the value false that speech or thought is represented indirectly, e.g. by use of a marked verbal aspect.

*Member of* model.attributable

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author biblScope cit citedRange corr del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled sp speaker stage street term textLang title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* actor camera caption castList epilogue performance prologue role roleDesc set sound tech view

*figures:* cell figDesc figure

*header:* change distributor edition extent geoDecl handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<!-- in the header --><editorialDecl>
  <quotation  marks="all"/>
</editorialDecl>
<!-- ... -->
<p>
  <said>"Our minstrel here will warm the old man's heart with song, dazzle him
with jewels and
    gold"</said>, a troublemaker simpered. <said>"He'll trample on the Duke's
camellias, spill
    his wine, and blunt his sword, and say his name begins with X, and in the
end the Duke
    will say, <said>'Take Saralinda, with my blessing, O lordly Prince of Rags
and Tags, O
        rider of the sun!'</said>"</said>
</p>
```

*Example*

```
<p>
  <said  aloud="true"  rend="pre(“) post(”)">Hmmm</said>, said a small voice in
```

```
his ear.
<said  aloud="true"  rend="pre(“) post(”)">Difficult. Very difficult. Plenty of
courage, I see.
    Not a bad mind either. there's talent, oh my goodness, yes — and a nice
thirst to prove
    yourself, now that's interesting. … So where shall I put you?</said>
</p>
<p>Harry gripped the edges of the stool and thought,
<said  aloud="false"  rend="italic">Not
    Slytherin, not Slytherin</said>.</p>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element said
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.directed.attributes,
   att.ascribed.attributes,
   att.cmc.attributes,
   attribute aloud { text }?,
   attribute direct { text }?,
   macro.specialPara}
```

---

## <salute>

(salutation) contains a salutation or greeting prefixed to a foreword, dedicatory epistle, or other division of a text, or the salutation in the closing of a letter, preface, etc. [*4.2.2. Openers and Closers*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.written
  - *@hand*

*Member of* model.divWrapper

*Contained by*

*cmc:* post

*core:* lg list

*drama:* castList epilogue performance prologue

*figures:* figure table

*textstructure:* body closer div div1 div2 div3 div4 div5 div6 div7 front group opener

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<salute>To all courteous mindes, that will voutchsafe the readinge.</salute>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element salute
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.written.attributes,
   macro.paraContent}
```

---

## <samplingDecl>

(sampling declaration) contains a prose description of the rationale and methods used in selecting texts, or parts of a text, for inclusion in the resource. [*2.3.2. The Sampling Declaration 2.3. The Encoding Description 16.3.2. Declarable Elements*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.declarable
  - *@default*

*Member of* model.encodingDescPart

*Contained by*

*header:* encodingDesc

*May contain*

*core:* p

*linking:* ab

*Note* This element records all information about systematic inclusion or omission of portions of the text, whether a reflection of sampling procedures in the pure sense or of systematic omission of material deemed either too difficult to transcribe or not of sufficient interest.

*Example*

```
<samplingDecl>
  <p>Samples of up to 2000 words taken at random from the beginning, middle, or
end of each
    text identified as relevant by respondents.</p>
</samplingDecl>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:samplingDecl"/> </sch:pattern>

*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element samplingDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   model.pLike+
}
```

---

**<schemaRef>** (schema reference) describes or points to a related customization or schema file. [*2.3.10. The Schema Specification*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

• att.resourced
– *@url*

• att.typed
– *@type*
– *@subtype*

`@key` the identifier used for the customization or schema.
*Status* Optional
*Datatype* teidata.xmlName

*Member of* model.encodingDescPart
*Contained by*
*header:* encodingDesc
*May contain*
*core:* desc
*Example*

```
<schemaRef type="interchangeODD"
  url="http://www.tei-c.org/release/xml/tei/custom/odd/tei_lite.odd"/>
<schemaRef type="interchangeRNG"
  url="http://www.tei-c.org/release/xml/tei/custom/odd/tei_lite.rng"/>
<schemaRef type="projectODD"
  url="file:///schema/project.odd"/>
```

*Content model*

```
<content>
 <classRef key="model.descLike"
  minOccurs="0"/>
</content>
```

*Schema Declaration*

```
element schemaRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.resourced.attributes,
   att.typed.attributes,
   attribute key { text }?,
   model.descLike?
}
```

---

## <schemaSpec>

(schema specification) generates a TEI-conformant schema and documentation for it. [*2.3. The Encoding Description 23.2. Modules and Schemas 24.5.1. Making a Unified ODD*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.docStatus
  - *@status*

- att.identified
  - *@ident*
  - *@predeclare*
  - *@module*
  - att.combinable
    - *@mode*
  - att.deprecated
    - *@validUntil*
- att.namespaceable
  - *@ns*

`@start` specifies entry points to the schema, i.e. which patterns may be used as the root of documents conforming to it.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.name separated by whitespace
*Default* TEI

`@prefix` specifies a default prefix which will be prepended to all patterns relating to TEI elements, unless otherwise stated.
*Status* Optional
*Datatype* 0–1 occurrences of teidata.xmlName separated by whitespace
*Note* Use of this attribute allows an external schema which has an element with the same local name as a TEI element to be mixed in.

`@targetLang` (target language) specifies which language to use when creating the objects in a schema if names for elements or attributes are available in more than one language.
*Status* Optional
*Datatype* teidata.language

`@docLang` (documentation language) specifies which languages to use when creating documentation if the description for an element, attribute, class or macro is available in more than one language.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.language separated by whitespace
*Note* Although multiple languages may be specified, the current TEI ODD processor is only capable of handling one.

`@defaultExceptions` (default namespace exclusions) provides a list of namespaces and/or prefixed element names to be excluded by default from anyName in RELAX NG schemas.
*Status* Recommended
*Datatype* 1–∞ occurrences of teidata.namespaceOrName separated by whitespace
*Default* http://www.tei-c.org/ns/1.0 teix:egXML
*Note* Specifies defaults for the *except* attribute of <anyElement>. RELAX NG validation (unless DTD Compatibility Mode is turned off) requires that any elements which may take an *xml:id* be excluded from the content of `anyName` name classes. For normal TEI purposes, this includes anything in the TEI namespace and the <egXML> element. The *defaultExceptions* attribute sets the global defaults for these exclusions. Element prefixes must be declared using a namespace declaration on the parent or an ancestor element.

*Member of* model.encodingDescPart model.frontPart
*Contained by*

*header:* encodingDesc

*textstructure:* back body div div1 div2 div3 div4 div5 div6 div7 front

*May contain*

*core:* desc gloss

*tagdocs:* classRef classSpec constraintDecl constraintSpec dataRef dataSpec elementRef elementSpec equiv listRef macroRef macroSpec moduleRef moduleSpec outputRendition specGrp specGrpRef

*Note* A <schemaSpec> combines references to modules, individual element or macro declarations, and specification groups together to form a unified schema.

The processing of the <schemaSpec> element must resolve any conflicts amongst the declarations it either contains or refers to. Different ODD processors may generate schemas and documentation using different concrete syntaxes.

The source may be specified (on the *source* attribute) in the form of a private URI, for which the recommended format is `tei:x.y.z`, where `x.y.z` indicates the version number, e.g. `tei:1.5.1` for 1.5.1 release of TEI P5 or (as a special case) `tei:current` for whatever is the latest release. The source indicated must provide a set of TEI-conformant specifications in a form directly usable by an ODD processor. By default, this will be the location of the current release of the TEI Guidelines.

*Example*

```
<schemaSpec prefix="TEI_" ident="testsvg"
  start="TEI svg">
  <moduleRef source="tei:current"
    key="header"/>
  <moduleRef key="core"/>
  <moduleRef key="drama"/>
  <moduleRef url="svg11.rng"/>
</schemaSpec>
```

This schema combines elements from the current versions of the drama and core modules, the current release version of the header module, and elements from an existing RELAX NG schema available from the URL indicated.

*Example*

```
<schemaSpec ident="Bare-plus"
  source="tei_bare.compiled.odd" start="TEI">
  <moduleRef key="tei"/>
  <moduleRef key="header"/>
  <elementRef key="q" source="tei:3.0.0"/>
  <moduleRef key="textstructure"/>
</schemaSpec>
```

This schema starts with a pre-existing customization of the TEI called TEI Bare (a customization with a minimal selection of elements). The <q> element is not available in TEI Bare, but it can be brought back. In this case, we will get the version defined in TEI P5 release 3.0.0

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.identEquiv"/>
   <classRef key="model.descLike"/>
  </alternate>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="constraintDecl"/>
```

```
   <classRef key="model.oddRef"/>
   <classRef key="model.oddDecl"/>
   <elementRef key="listRef"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element schemaSpec
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.docStatus.attributes,
   att.identified.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   att.namespaceable.attributes,
   attribute start { list { + } }?,
   attribute prefix { ? }?,
   attribute targetLang { text }?,
   attribute docLang { list { + } }?,
   attribute defaultExceptions { list { + } }?,
   (
      ( model.identEquiv | model.descLike )*,
      ( constraintDecl | model.oddRef | model.oddDecl | listRef )*
   )
}
```

---

## <scriptDesc>

<scriptDesc> (script description) contains a description of the scripts used in a manuscript or other object.
[*11.7.2.1. Writing*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs

    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.physDescPart
*Contained by*
*msdescription:* physDesc
*May contain*
*core:* p
*header:* scriptNote
*linking:* ab
*msdescription:* summary
*Example*

```
<scriptDesc>
  <p/>
</scriptDesc>
```

*Example*

```
<scriptDesc>
  <summary>Contains two distinct styles of scripts </summary>
  <scriptNote  xml:id="style-1">.</scriptNote>
  <scriptNote  xml:id="style-2">.</scriptNote>
</scriptDesc>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="summary" minOccurs="0"/>
```

```
   <elementRef key="scriptNote"
    minOccurs="1" maxOccurs="unbounded"/>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element scriptDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.pLike+ | ( summary?, scriptNote+ ) )
}
```

---

**<scriptNote>** describes a particular script distinguished within the description of a manuscript or similar resource. [*11.7.2. Writing, Decoration, and Other Notations*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.handFeatures
  - *@scribe*
  - *@scribeRef*
  - *@script*
  - *@scriptRef*
  - *@medium*

  - att.scope
    * *@scope*

*Contained by msdescription:* scriptDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<scriptNote scope="sole"/>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element scriptNote
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.handFeatures.attributes,
   att.scope.attributes,
   macro.specialPara}
```

---

**<scriptStmt>** (script statement) contains a citation giving details of the script used for a spoken text. [*8.2. Documenting the Source of Transcribed Speech 2.2.7. The Source Description 16.3.2. Declarable Elements*]

*Module* spoken — Transcriptions of Speech

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  – att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  – att.global.responsibility
    * *@cert*
    * *@resp*
  – att.global.source
    * *@source*
- att.declarable
  – *@default*

*Member of* model.sourceDescPart
*Contained by*
*header:* sourceDesc
*May contain*
*core:* bibl biblStruct listBibl p
*header:* biblFull
*linking:* ab
*msdescription:* msDesc
*Example*

```
<scriptStmt>
  <bibl>
    <author>Craig Warner</author>
    <title>Strangers on a Train</title>
    <title  type="sub">Based on the novel by Patricia Highsmith</title>
    <edition>French's acting edition</edition>
    <idno  type="ISBN">978 0 573 01972 2</idno>
    <publisher>Samuel French Ltd</publisher>
  </bibl>
</scriptStmt>
```

*Schematron* <sch:pattern is-a=”declarable”> <sch:param name=”tde” value=”tei:scriptStmt”/> </sch:pattern>
*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <classRef key="model.biblLike"/>
 </alternate>
```

```
</content>
```

*Schema Declaration*

```
element scriptStmt
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   ( model.pLike+ | model.biblLike )
}
```

---

**<seal>** (seal) contains a description of one seal or similar applied to the object described. [*11.7.3.2. Seals*]

*Module* msdescription — Manuscript Description

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*

  * *@rendition*
  - att.global.responsibility
  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.datable
  - *@period*
  - att.datable.custom
  * *@when-custom*
  * *@notBefore-custom*
  * *@notAfter-custom*
  * *@from-custom*
  * *@to-custom*
  * *@datingPoint*
  * *@datingMethod*
  - att.datable.iso
  * *@when-iso*
  * *@notBefore-iso*
  * *@notAfter-iso*
  * *@from-iso*
  * *@to-iso*
  - att.datable.w3c
  * *@when*
  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*
- att.typed
  - *@type*
  - *@subtype*

`@contemporary` (contemporary) specifies whether or not the seal is contemporary with the item to which it is affixed
*Status* Optional
*Datatype* teidata.xTruthValue

*Contained by:* msdescription: sealDesc
*May contain*
*core:* p
*linking:* ab
*msdescription:* decoNote
*Example*

```
<seal  n="2"  type="pendant"
  subtype="cauda_duplex">
  <p>The seal of <name>Jens Olufsen</name> in black wax.
     (<ref>DAS 1061</ref>). Legend: <q>S IOHANNES OLAVI</q>.
     Parchment tag on which is written: <q>Woldorp Iohanne G</q>.</p>
</seal>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.pLike"/>
  <elementRef key="decoNote"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element seal
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.typed.attributes,
   attribute contemporary { text }?,
   ( model.pLike | decoNote )+
}
```

---

## <sealDesc>

<sealDesc> (seal description) describes the seals or similar items related to the object described, either as a series of paragraphs or as a series of <seal> elements. [*11.7.3.2. Seals*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
- att.global.analytic
  - *@ana*

- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.physDescPart
*Contained by*
*msdescription:* physDesc
*May contain*
*core:* p
*linking:* ab
*msdescription:* condition decoNote seal summary
*Example*

```
<sealDesc>
  <seal  type="pendant"  contemporary="true">
    <p>Green wax vertical oval seal attached at base.</p>
  </seal>
</sealDesc>
```

*Example*

```
<sealDesc>
  <p>Parchment strip for seal in place; seal missing.</p>
</sealDesc>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
```

```
  <sequence>
   <elementRef key="summary" minOccurs="0"/>
   <alternate minOccurs="1"
    maxOccurs="unbounded">
    <elementRef key="decoNote"/>
    <elementRef key="seal"/>
    <elementRef key="condition"/>
   </alternate>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element sealDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.pLike+ | ( summary?, ( decoNote | seal | condition )+ ) )
}
```

---

**<secFol>** (second folio) marks the word or words taken from a fixed point in a codex (typically the beginning of the second leaf) in order to provide a unique identifier for it. [*11.3.7. Catchwords, Signatures, Secundo Folio*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*

    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<secFol>(con-)versio morum</secFol>
```

*Schematron* <sch:rule context="tei:secFol"> <sch:assert test="ancestor::tei:msDesc or ancestor::tei:egXML">The <sch:name/> element should not be used outside of msDesc.</sch:assert> </sch:rule>

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element secFol
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

**<secl>** (secluded text) Secluded. Marks text present in the source which the editor believes to be genuine but out of its original place (which is unknown). [*12.3.1.7. Text Omitted from or Supplied in the Transcription*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*
  - att.ranging
    - *@atLeast*
    - *@atMost*

* *@min*
* *@max*
* *@confidence*

- att.editLike
  - *@evidence*
  - *@instant*

`@reason` one or more words indicating why this text has been secluded, e.g. *interpolated* etc.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace

*Member of* model.pPart.transcriptional
*Contained by*
*analysis:* cl pc phr s w
*cmc:* post
*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castItem role roleDesc sound tech view
*figures:* cell
*header:* change distributor edition extent geoDecl handNote licence scriptNote
*linking:* ab seg
*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark
*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname
*spoken:* u writing
*tagdocs:* eg
*textcrit:* lem rdg wit witDetail
*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer
*transcr:* am damage fw metamark mod restore retrace secl supplied surplus
*verse:* rhyme
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound tech view
*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<rdg  source="#Pescani">
  <secl>
    <l  n="15"  xml:id="l15">Alphesiboea suos ulta est pro coniuge fratres,</l>
    <l  n="16"  xml:id="l16">sanguinis et cari vincula rupit amor.</l>
  </secl>
</rdg>
<note>secl. Pescani</note>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element secl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   attribute reason { list { + } }?,
   macro.paraContent}
```

**<seg>** (arbitrary segment) represents any segmentation of text below the 'chunk' level. [*17.3. Blocks, Segments, and Anchors 6.2. Components of the Verse Line 7.2.5. Speech Contents*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.notated
  - *@notation*
- att.segLike
  - *@function*
  - att.datcat
    - *@datcat*

    - * *@valueDatcat*
    - * *@targetDatcat*
    - att.fragmentable
    - * *@part*
    - att.metrical
    - * *@met*
    - * *@real*
    - * *@rhyme*
  - att.typed
    - *@type*
    - *@subtype*
  - att.written
    - *@hand*

*Member of* model.choicePart model.linePart model.segLike model.standOffPart

*Contained by*

*analysis:* cl m phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell notatedMusic

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media

mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The <seg> element may be used at the encoder's discretion to mark any segments of the text of interest for processing. One use of the element is to mark text features for which no appropriate markup is otherwise defined. Another use is to provide an identifier for some segment which is to be pointed at by some other element—i.e. to provide a target, or a part of a target, for a <ptr> or other similar element.

*Example*

```
<seg>When are you leaving?</seg>
<seg>Tomorrow.</seg>
```

*Example*

```
<s>
  <seg  rend="caps"  type="initial-cap">So father's only</seg> glory was the
ballfield.
</s>
```

*Example*

```
<seg  type="preamble">
  <seg>Sigmund, <seg  type="patronym">the son of Volsung</seg>, was a king in
Frankish country.</seg>
  <seg>Sinfiotli was the eldest of his sons ...</seg>
  <seg>Borghild, Sigmund's wife, had a brother ... </seg>
</seg>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
```

```
</content>
```

*Schema Declaration*

```
element seg
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.notated.attributes,
   att.segLike.attributes,
   att.datcat.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   att.written.attributes,
   macro.paraContent}
```

---

## <segmentation>

**<segmentation>** (segmentation) describes the principles according to which the text has been segmented, for example into sentences, tone-units, graphemic strata, etc. [*2.3.3. The Editorial Practices Declaration 16.3.2. Declarable Elements*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.declarable
  - *@default*

*Member of* model.editorialDeclPart

*Contained by*

*header:* editorialDecl

*May contain*

*core:* p

*linking:* ab

*Example*

```
<segmentation>
  <p>
    <gi>s</gi> elements mark orthographic sentences and are numbered
sequentially within
     their parent <gi>div</gi> element </p>
</segmentation>
```

*Example*

```
<p>
  <gi>seg</gi> elements are used to mark functional constituents of various
types within each
<gi>s</gi>; the typology used is defined by a <gi>taxonomy</gi> element in the
corpus
header <gi>classDecl</gi>
</p>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:segmentation"/> </sch:pattern>

*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element segmentation
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   model.pLike+
}
```

---

**<sense>** groups together all information relating to one word sense in a dictionary entry, for example definitions, examples, and translation equivalents. [*10.2. The Structure of Dictionary Entries*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*

* *@resp*

– att.global.source
* *@source*

• att.lexicographic
– *@expand*
– *@split*
– *@value*
– *@location*
– *@mergedIn*
– *@opt*

– att.datcat
* *@datcat*
* *@valueDatcat*
* *@targetDatcat*

– att.lexicographic.normalized
* *@norm*
* *@orig*

`@level` gives the nesting depth of this sense.
*Status* Optional
*Datatype* teidata.count

*Member of* model.entryPart
*Contained by*
*core:* cit
*dictionaries:* dictScrap entry entryFree hom re sense
*namesdates:* nym
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q ref reg rs ruby sic soCalled term time title unclear unit
*dictionaries:* def dictScrap entry etym form gramGrp lang lbl oRef pRef re sense usg xr
*figures:* figure formula notatedMusic
*gaiji:* g
*header:* biblFull idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May contain character data mixed with any other elements defined in the dictionary tag set.

*Example*

```
<sense  n="2">
  <usg  type="time">Vx.</usg>
  <def>Vaillance, bravoure (spécial., au combat)</def>
  <cit  type="example">
    <quote>La valeur n'attend pas le nombre des années</quote>
    <bibl>
      <author>Corneille</author>
    </bibl>
  </cit>
</sense>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <elementRef key="sense"/>
  <classRef key="model.entryPart.top"/>
  <classRef key="model.phrase"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element sense
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   attribute level { text }?,
   (
      text
    | model.gLike    | sense    | model.entryPart.top    | model.phrase    | model.global
}
```

**<sequence>** indicates that the constructs referenced by its children form a sequence. [*23.5.1. Defining Content Models*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.repeatable
  - *@minOccurs*
  - *@maxOccurs*

`@preserveOrder` if false, indicates that component elements of a sequence may occur in any order.
*Status* Optional
*Datatype* teidata.truthValue

*Member of* model.contentPart

*Contained by*

*tagdocs:* alternate content sequence

*May contain*

*tagdocs:* alternate anyElement classRef dataRef elementRef empty macroRef sequence textNode

*Example*

```
<content>
  <sequence>
    <alternate>
      <elementRef  key="name"/>
      <elementRef  key="persName"/>
    </alternate>
    <elementRef  key="placeName"  minOccurs="0"
      maxOccurs="5"/>
  </sequence>
</content>
```

This example content model matches a sequence consisting of either a <name> or a <persName> followed by nothing, or by a sequence of up to five <placeName> elements.

*Schematron* <sch:rule context="tei:sequence"> <sch:assert test="count(*) gt 1">The sequence element must have at least two child elements</sch:assert> </sch:rule>

*Content model*

```
<content>
 <classRef key="model.contentPart"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element sequence
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.repeatable.attributes,
   attribute preserveOrder { text }?,
   model.contentPart+
}
```

---

## <series>

<series> (series information) contains information about the series in which a book or other bibliographic item has appeared. [*3.12.2.1. Analytic, Monographic, and Series Levels*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.biblPart
*Contained by*
*core:* bibl biblStruct
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* biblScope cb editor ellipsis gap gb index lb milestone note noteGrp pb ptr ref respStmt textLang title
*figures:* figure notatedMusic
*gaiji:* g
*header:* availability idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* listRef
*textcrit:* app witDetail
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
character data
*Example*

```
<series xml:lang="de">
  <title level="s">Halbgraue Reihe zur Historischen Fachinformatik</title>
  <respStmt>
    <resp>Herausgegeben von</resp>
    <name type="person">Manfred Thaller</name>
    <name type="org">Max-Planck-Institut für Geschichte</name>
  </respStmt>
  <title level="s">Serie A: Historische Quellenkunden</title>
  <biblScope>Band 11</biblScope>
</series>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <elementRef key="title"/>
  <classRef key="model.ptrLike"/>
  <elementRef key="editor"/>
  <elementRef key="respStmt"/>
  <elementRef key="biblScope"/>
  <elementRef key="idno"/>
  <elementRef key="textLang"/>
  <classRef key="model.global"/>
  <elementRef key="availability"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element series
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   (
      text
    | model.gLike    | title    | model.ptrLike    | editor    | respStmt    | biblScope
}
```

---

**<seriesStmt>** (series statement) groups information about the series, if any, to which a publication belongs.
[*2.2.5. The Series Statement 2.2. The File Description*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*
- att.declarable
  - *@default*

*Contained by:* biblFull fileDesc
*May contain*
*core:* biblScope editor p respStmt title
*header:* idno
*linking:* ab
*Example*

```
<seriesStmt>
  <title>Machine-Readable Texts for the Study of Indian Literature</title>
  <respStmt>
    <resp>ed. by</resp>
    <name>Jan Gonda</name>
  </respStmt>
  <biblScope  unit="volume">1.2</biblScope>
```

```
      <idno  type="ISSN">0 345 6789</idno>
    </seriesStmt>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:seriesStmt"/> </sch:pattern>
*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="title" minOccurs="1"
    maxOccurs="unbounded"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="editor"/>
    <elementRef key="respStmt"/>
   </alternate>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="idno"/>
    <elementRef key="biblScope"/>
   </alternate>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element seriesStmt
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   ( model.pLike+ | ( title+, ( editor | respStmt )*, ( idno | biblScope )* ) )
}
```

---

**<set>** (setting) contains a description of the setting, time, locale, appearance, etc., of the action of a play, typically found in the front matter of a printed performance text (not a stage direction). [*7.1. Front and Back Matter* ]

*Module* drama — Performance Texts
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.frontPart.drama

*Contained by*

*textcrit:* lem rdg

*textstructure:* back front

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*cmc:* post

*core:* bibl biblStruct cb cit desc ellipsis gap gb head index l label lb lg list listBibl milestone note noteGrp p pb q quote said sp stage

*dictionaries:* entry entryFree superEntry

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure notatedMusic table

*header:* biblFull

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* msDesc

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* classSpec constraintSpec dataSpec eg egXML elementSpec macroSpec moduleSpec outputRendition specGrp specGrpRef

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* Contains paragraphs or phrase level tags.

This element should not be used outside the front or back matter; for similar contextual descriptions within the body of the text, use the <stage> element.

*Example*

```
<set>
  <p>The action takes place on February 7th between the hours of noon and six
in the
    afternoon, close to the Trenartha Tin Plate Works, on the borders of
England and Wales,
    where a strike has been in progress throughout the winter.</p>
</set>
```

*Example*

```
<set>
  <head>SCENE</head>
  <p>A Sub-Post Office on a late autumn evening</p>
</set>
```

*Example*

```
<front>
<!-- <titlePage>, <div type="Dedication">, etc. -->
  <set>
    <list  type="gloss">
      <label>TIME</label>
      <item>1907</item>
      <label>PLACE</label>
      <item>East Coast village in England</item>
    </list>
  </set>
</front>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.headLike"/>
   <classRef key="model.global"/>
  </alternate>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.common"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
```

```
</content>
```

*Schema Declaration*

```
element set
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( ( model.headLike | model.global )*, ( ( model.common, model.global* )* ) )
}
```

---

**<setting>** describes one particular setting in which a language interaction takes place. [*16.2.3. The Setting Description*]

*Module* corpus — Language Corpora

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*

    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.ascribed
  - *@who*

*Contained by:* *core:* settingDesc

*May contain*

*core:* date name p time

*corpus:* activity locale

*linking:* ab

*namesdates:* orgName persName placeName

*Note* If the *who* attribute is not supplied, the setting is assumed to be that of all participants in the language interaction.

*Example*

```
<setting>
  <placeName>New York City, US</placeName>
  <date>1989</date>
  <locale>on a park bench</locale>
  <activity>feeding birds</activity>
</setting>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.nameLike.agent"/>
   <classRef key="model.dateLike"/>
   <classRef key="model.settingPart"/>
  </alternate>
 </alternate>
</content>
```

*Schema Declaration*

```
element setting
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
```

```
   att.ascribed.attributes,
   (
      model.pLike+
    | ( model.nameLike.agent | model.dateLike | model.settingPart )*
   )
}
```

---

**<settingDesc>** (setting description) describes the setting or settings within which a language interaction takes place, or other places otherwise referred to in a text, edition, or metadata. [*16.2. Contextual Information 2.4. The Profile Description*]

*Module* corpus — Language Corpora

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable

– *@default*

*Member of* model.profileDescPart

*Contained by*

*header:* profileDesc

*May contain*

*core:* p

*corpus:* setting

*linking:* ab

*namesdates:* listPlace place

*Note* May contain a prose description organized as paragraphs, or a series of <setting> elements. If used to record not settings of language interactions, but other places mentioned in the text, then <place> optionally grouped by <listPlace> inside <standOff> should be preferred.

*Example*

```
<settingDesc>
  <p>Texts recorded in the
     Canadian Parliament building in Ottawa, between April and November 1988
</p>
</settingDesc>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:settingDesc"/> </sch:pattern>

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <elementRef key="setting"/>
   <classRef key="model.placeLike"/>
   <elementRef key="listPlace"/>
  </alternate>
 </alternate>
</content>
```

*Schema Declaration*

```
element settingDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   ( model.pLike+ | ( setting | model.placeLike | listPlace )+ )
}
```

**<settlement>** (settlement) contains the name of a settlement such as a city, town, or village identified as a single geo-political or administrative unit. [*14.2.3. Place Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*
  - att.datable.custom
    - *@when-custom*
    - *@notBefore-custom*
    - *@notAfter-custom*
    - *@from-custom*

    - *@to-custom*
    - *@datingPoint*
    - *@datingMethod*
  - att.datable.iso
    - *@when-iso*
    - *@notBefore-iso*
    - *@notAfter-iso*
    - *@from-iso*
    - *@to-iso*
  - att.datable.w3c
    - *@when*
    - *@notBefore*
    - *@notAfter*
    - *@from*
    - *@to*
- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
    - *@key*
    - *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.placeNamePart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage unitDef

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions altIdentifier catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material

msIdentifier musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectIdentifier objectName occupation offset org orgName persName persPronouns place placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<placeName>
  <settlement  type="town">Glasgow</settlement>
  <region>Scotland</region>
</placeName>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
```

```
</content>
```

*Schema Declaration*

```
element settlement
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<sex>** (sex) specifies the sex of an organism. [*14.3.2.1. Personal Characteristics*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.editLike
  - *@evidence*
  - *@instant*
- att.typed
  - *@type*
  - *@subtype*

`@value` supplies a coded value for sex.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.sex separated by whitespace

*Note* Values for this attribute may be locally defined by a project, or they may refer to an external standard.

*Member of* model.persStateLike

*Contained by*

*namesdates:* person personGrp persona

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* As with other culturally-constructed traits such as age and gender, the way in which this concept is described in different cultural contexts varies. The normalizing attributes are provided only as an optional means of simplifying that variety for purposes of interoperability or project-internal taxonomies for consistency, and should not be used where that is inappropriate or unhelpful. The content of the element may be used to describe the intended concept in more detail.

*Example*

```
<sex value="F">female</sex>
```

*Example*

```
<sex value="I">Intersex</sex>
```

*Example*

```
<sex value="TG F">Female (TransWoman)</sex>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
```

```
</content>
```

*Schema Declaration*

```
element sex
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.typed.attributes,
   attribute value { list { + } }?,
   macro.phraseSeq}
```

---

**<shift>** (shift) marks the point at which some paralinguistic feature of a series of utterances by any one speaker changes. [*8.3.6. Shifts*]

*Module* spoken — Transcriptions of Speech

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

  * *@select*
  - att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
  - att.global.responsibility
  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.ascribed
  - *@who*
- att.cmc
  - *@generatedBy*

`@feature` a paralinguistic feature.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **tempo** speed of utterance.
**loud** loudness.
**pitch** pitch range.
**tension** tension or stress pattern.
**rhythm** rhythmic qualities.
**voice** voice quality.

`@new` specifies the new state of the paralinguistic feature specified.
*Status* Recommended
*Datatype* teidata.enumerated
*Default* normal
*Note* Some possible values for this attribute are provided in section *8.4.2. Synchronization and Overlap*. The special value normal should be used to indicate that the feature concerned ceases to be remarkable at this point. In earlier versions of these Guidelines, a null value for this attribute was understood to have the same effect: this practice is now deprecated and will be removed at a future release.

*Member of* model.global.spoken
*Contained by*
*analysis:* cl m phr s span w
*cmc:* post
*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit
*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose
*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view
*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* annotationBlock u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain* Empty element

*Example*

```
<u who="#LB">
  <shift feature="loud" new="f"/>Elizabeth
</u>
<u who="#EB">Yes</u>
<u who="#LB">
  <shift feature="loud" new="normal"/>Come and try this
<pause/>
  <shift feature="loud" new="ff"/>come on
</u>
<!-- ... -->
<listPerson type="speakers">
  <person xml:id="LB"/>
  <person xml:id="EB"/>
</listPerson>
```

The word ‘Elizabeth’ is spoken loudly, the words ‘Yes’ and ‘Come and try this’ with normal volume, and the words ‘come on’ very loudly.

*Schematron* <sch:rule context=”tei:shift”> <sch:assert test=”@new” role=”warning”> The @new attribute should always be supplied; use the special value ”normal” to indicate that the feature concerned ceases to be remarkable at this point. </sch:assert> </sch:rule>

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element shift
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
```

```
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.attributes,
   att.cmc.attributes,
   attribute feature
   {
      "tempo" | "loud" | "pitch" | "tension" | "rhythm" | "voice"
   }?,
   attribute new { text }?,
   empty
}
```

---

**<sic>** (Latin for thus or so) contains text reproduced although apparently incorrect or inaccurate. [*3.5.1. Apparent Errors*]

*Module* core — Elements Available in All TEI Documents

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*

- att.global.source
  - *@source*
- att.cmc
  - *@generatedBy*

*Member of* model.choicePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
for his nose was as sharp as
a pen, and <sic>a Table</sic> of green fields.
```

*Example* If all that is desired is to call attention to the apparent problem in the copy text, <sic> may be used alone:

```
I don't know, Juan. It's so far in the past now
— how <sic>we can</sic> prove or disprove anyone's theories?
```

*Example* It is also possible, using the <choice> and <corr> elements, to provide a corrected reading:

```
I don't know, Juan. It's so far in the past now
— how <choice>
  <sic>we can</sic>
  <corr>can we</corr>
</choice> prove or disprove anyone's theories?
```

*Example*

```
for his nose was as sharp as
a pen, and <choice>
  <sic>a Table</sic>
  <corr>a' babbld</corr>
</choice> of green fields.
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element sic
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
```

```
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   macro.paraContent}
```

---

**<signatures>** (signatures) contains discussion of the leaf or quire signatures found within a codex or similar object. [*11.3.7. Catchwords, Signatures, Secundo Folio*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<signatures>Quire and leaf signatures in letters, [b]-v, and roman
numerals; those in quires 10 (1) and 17 (s) in red ink and different
from others; every third quire also signed with red crayon in arabic
numerals in the center lower margin of the first leaf recto: "2" for
quire 4 (f. 19), "3" for quire 7 (f. 43); "4," barely visible, for
quire 10 (f. 65), "5," in a later hand, for quire 13 (f. 89), "6," in
a later hand, for quire 16 (f. 113).</signatures>
```

*Schematron* <sch:rule context="tei:signatures"> <sch:assert test="ancestor::tei:msDesc or ancestor::tei:egXML">The <sch:name/> element should not be used outside of msDesc.</sch:assert> </sch:rule>

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element signatures
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.specialPara}
```

---

## <signed>

(signature) contains the closing salutation, etc., appended to a foreword, dedicatory epistle, or other division of a text. [*4.2.2. Openers and Closers*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.written
  - *@hand*

*Member of* model.divBottomPart model.divTopPart

*Contained by*

*cmc:* post

*core:* lg list

*drama:* castList epilogue performance prologue

*figures:* figure table

*textstructure:* back body closer div div1 div2 div3 div4 div5 div6 div7 front group opener postscript

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<signed>Thine to command <name>Humph. Moseley</name>
</signed>
```

*Example*

```
<closer>
  <signed>Sign'd and Seal'd,
  <list>
     <item>John Bull,</item>
     <item>Nic. Frog.</item>
   </list>
  </signed>
</closer>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element signed
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
```

```
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.written.attributes,
   macro.paraContent}
```

---

## <soCalled>

(so called) contains a word or phrase for which the author or narrator indicates a disclaiming of responsibility, for example by the use of scare quotes or italics. [*3.3.3. Quotation*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*

*Member of* model.emphLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
To edge his way along
the crowded paths of life, warning all human sympathy to keep its distance, was
what the
knowing ones call <soCalled>nuts</soCalled> to Scrooge.
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element soCalled
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   macro.phraseSeq}
```

---

## <socecStatus>

(socio-economic status) contains an informal description of a person's perceived social or economic status. [*16.2.2. The Participant Description*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change

    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*

  * *@to*

- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*
  - att.canonical
  * *@key*
  * *@ref*
- att.typed
  - ~~type~~
  - @subtype

`@type` characterizes the element in some sense, using any convenient classification scheme or typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **atBirth**
**atDeath**
**dependent**
**inherited**
**independent**

`@scheme` identifies the classification system or taxonomy in use, for example by pointing to a locally-defined <taxonomy> element or by supplying a URI for an externally-defined system.
*Status* Optional
*Datatype* teidata.pointer

`@code` identifies a status code defined within the classification system or taxonomy defined by the *scheme* attribute.
*Status* Optional
*Datatype* teidata.pointer

*Member of* model.persStateLike
*Contained by*
*namesdates:* person personGrp persona
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit
*dictionaries:* lang oRef pRef
*figures:* figure formula notatedMusic
*gaiji:* g
*header:* idno
*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The content of this element may be used as an alternative to the more formal specification made possible by its attributes; it may also be used to supplement the formal specification with commentary or clarification.

*Example*

```
<socecStatus  scheme="#rg"  code="#ab1"/>
```

*Example*

```
<socecStatus>Status AB1 in the RG Classification scheme</socecStatus>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element socecStatus
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   attribute scheme { text }?,
```

```
attribute code { text }?,
macro.phraseSeq}
```

---

**<sound>** (sound) describes a sound effect or musical sequence specified within a screen play or radio script.
[*7.3.1. Technical Information 7.3. Other Types of Performance Text*]

*Module* drama — Performance Texts

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - ~~type~~
  - @subtype

`@type` categorizes the sound in some respect, e.g. as music, special effect, etc.
*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

`@discrete` indicates whether the sound overlaps the surrounding speeches or interrupts them.

*Status* Optional

*Datatype* teidata.xTruthValue

*Note* The value true indicates that the sound is heard between the surrounding speeches; the value false indicates that the sound overlaps one or more of the surrounding speeches.

*Member of* model.stageLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l lg meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound spGrp tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink

objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* A specialized form of stage direction.

*Example*

```
<sp>
  <speaker>Benjy</speaker>
  <p>Now to business.</p>
</sp>
<sp>
  <speaker>Ford and Zaphod</speaker>
  <p>To business.</p>
</sp>
<sound  discrete="true">Glasses clink.</sound>
<sp>
  <speaker>Benjy</speaker>
  <p>I beg your pardon?</p>
</sp>
<sp>
  <speaker>Ford</speaker>
  <p>I'm sorry, I thought you were proposing a toast.</p>
</sp>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element sound
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   attribute discrete { text }?,
   macro.paraContent}
```

**<source>** (source) describes the original source for the information contained with a manuscript or object description. [*11.9.1.1. Record History*]

*Module* msdescription — Manuscript Description

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Contained by* msDescription: recordHist

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<source>Derived from <ref>Stanley (1960)</ref>
</source>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element source
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.specialPara}
```

**<sourceDesc>** (source description) describes the source(s) from which an electronic text was derived or generated, typically a bibliographic description in the case of a digitized text, or a phrase such as 'born digital' for a text which has no previous existence. [*2.2.7. The Source Description*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable
  - *@default*

*Contained by* *header:* biblFull fileDesc

*May contain*

*core:* bibl biblStruct list listBibl p

*figures:* table

*header:* biblFull

*linking:* ab

*msdescription:* msDesc
*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation
*spoken:* recordingStmt scriptStmt
*textcrit:* listApp listWit
*Example*

```
<sourceDesc>
  <bibl>
    <title  level="a">The Interesting story of the Children in the Wood</title>.
In
  <author>Victor E Neuberg</author>, <title>The Penny Histories</title>.
  <publisher>OUP</publisher>
    <date>1968</date>. </bibl>
</sourceDesc>
```

*Example*

```
<sourceDesc>
  <p>Born digital: no previous source exists.</p>
</sourceDesc>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:sourceDesc"/> </sch:pattern>
*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.biblLike"/>
   <classRef key="model.sourceDescPart"/>
   <classRef key="model.listLike"/>
  </alternate>
 </alternate>
</content>
```

*Schema Declaration*

```
element sourceDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   (
      model.pLike+
    | ( model.biblLike | model.sourceDescPart | model.listLike )+
   )
}
```

**<sourceDoc>** contains a transcription or other representation of a single source document potentially forming part of a *dossier génétique* or collection of sources. [*12.1. Digital Facsimiles 12.2.2. Embedded Transcription*]

*Module* transcr — Representation of Primary Sources

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declaring
  - *@decls*

*Member of* model.resource

*Contained by*

*core:* teiCorpus

*textstructure:* TEI

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* binaryObject cb ellipsis gap gb graphic index lb media milestone note noteGrp pb

*figures:* figure formula notatedMusic

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin surface surfaceGrp

*Note* This element may be used as an alternative to <facsimile> for TEI documents containing only page images, or for documents containing both images and transcriptions. Transcriptions may be provided within the <surface> elements making up a source document, in parallel with them as part of a <text> element, or in both places if the encoder wishes to distinguish these two modes of transcription.

*Example*

```
<sourceDoc>
  <surfaceGrp  n="leaf1">
    <surface  facs="page1.png">
      <zone>All the writing on page 1</zone>
    </surface>
    <surface>
      <graphic  url="page2-highRes.png"/>
      <graphic  url="page2-lowRes.png"/>
      <zone>
        <line>A line of writing on page 2</line>
        <line>Another line of writing on page 2</line>
      </zone>
    </surface>
  </surfaceGrp>
</sourceDoc>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.global"/>
  <classRef key="model.graphicLike"/>
  <elementRef key="surface"/>
  <elementRef key="surfaceGrp"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element sourceDoc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
```

```
   ( model.global | model.graphicLike | surface | surfaceGrp )+
}
```

---

**<sp>** (speech) contains an individual speech in a performance text, or a passage presented as such in a prose or verse text. [*3.13.2. Core Tags for Drama 3.13. Passages of Verse or Drama 7.2.2. Speeches and Speakers*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.ascribed.directed
  - *@toWhom*
  - att.ascribed
    - *@who*

- att.placement
  - *@place*
- att.written
  - *@hand*

*Member of* model.divPart

*Contained by*

*core:* item note q quote said stage textLang

*drama:* castList epilogue performance prologue set spGrp view

*figures:* cell figure

*header:* change handNote licence scriptNote

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*tagdocs:* specGrp

*textcrit:* lem rdg

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript

*transcr:* metamark

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* cb cit ellipsis gap gb index l lb lg list milestone note noteGrp p pb q quote said speaker stage

*drama:* camera caption move sound tech view

*figures:* figure notatedMusic table

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp timeline

*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* The *who* attribute on this element may be used either in addition to the <speaker> element or as an alternative.

*Example*

```
<sp>
  <speaker>The reverend Doctor Opimian</speaker>
  <p>I do not think I have named a single unpresentable fish.</p>
</sp>
<sp>
  <speaker>Mr Gryll</speaker>
  <p>Bream, Doctor: there is not much to be said for bream.</p>
</sp>
<sp>
  <speaker>The Reverend Doctor Opimian</speaker>
  <p>On the contrary, sir, I think there is much to be said for him. In the
first place [...]</p>
  <p>Fish, Miss Gryll — I could discourse to you on fish by the hour: but for
the present I
    will forbear [...]</p>
</sp>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.global"
    minOccurs="1" maxOccurs="1"/>
   <classRef key="model.stageLike"
    minOccurs="1" maxOccurs="1"/>
  </alternate>
  <sequence>
   <elementRef key="speaker"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <sequence minOccurs="1"
   maxOccurs="unbounded">
   <alternate>
    <elementRef key="lg"/>
    <classRef key="model.lLike"/>
    <classRef key="model.pLike"/>
    <classRef key="model.listLike"/>
    <classRef key="model.stageLike"/>
    <classRef key="model.attributable"/>
   </alternate>
   <alternate>
    <classRef key="model.global"
     minOccurs="0" maxOccurs="unbounded"/>
    <elementRef key="q"/>
   </alternate>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element sp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.directed.attributes,
   att.ascribed.attributes,
   att.placement.attributes,
   att.written.attributes,
   (
      ( model.global | model.stageLike )*,
      ( speaker, model.global* ),
      (
         (
            (
               lg             | model.lLike             | model.pLike         |     | model.
```

```
            ( model.global* | q )
         )+
      )
   )
}
```

---

**<space>** (space) indicates the location of a significant space in the text. [*12.4.1. Space*]

*Module* transcr — Representation of Primary Sources

*Attributes*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*
  - att.ranging
    - *@atLeast*
    - *@atMost*
    - *@min*
    - *@max*
    - *@confidence*
- att.typed
  - *@type*
  - *@subtype*
- att.global
  - @xml:id
  - @n
  - @xml:lang
  - @xml:base
  - @xml:space
- att.global.analytic
  - @ana
- att.global.change
  - @change
- att.global.facs
  - @facs
- att.global.linking
  - @corresp
  - @synch
  - @sameAs
  - @copyOf
  - @next
  - @prev

- @exclude
- @select

- att.global.rendition
  - @rend
  - @style
  - @rendition
- att.global.responsibility
  - ~~resp~~
  - @cert
- att.global.source
  - @source

`@resp` (responsible party) (responsible party) indicates the individual responsible for identifying and measuring the space.
*Derived from* att.global.responsibility
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

`@dim` (dimension) indicates whether the space is horizontal or vertical.
*Status* Recommended
*Datatype* teidata.enumerated
*Legal values are:* **horizontal** the space is horizontal.
**vertical** the space is vertical.
*Note* For irregular shapes in two dimensions, the value for this attribute should reflect the more important of the two dimensions. In conventional left-right scripts, a space with both vertical and horizontal components should be classed as vertical.

*Member of* model.global.edit
*Contained by*
*analysis:* cl m phr s span w
*cmc:* post
*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit
*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose
*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view
*figures:* cell figure table
*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor
*linking:* ab seg
*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark
*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName

occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*certainty:* certainty precision respons

*core:* desc

*Note* This element should be used wherever it is desired to record an unusual space in the source text, e.g. space left for a word to be filled in later, for later rubrication, etc. It is not intended to be used to mark normal inter-word space or the like.

*Example*

```
By god if wommen had writen storyes
As <space  quantity="7"  unit="minims"/> han within her oratoryes
```

*Example*

```
στρατηλάτ<space  quantity="1"  unit="chars"/>ου
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.descLike"/>
  <classRef key="model.certLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element space
{
   att.dimensions.attributes,
   att.ranging.attributes,
   att.typed.attributes,
   att.global.attribute.xmlid,
   att.global.attribute.n,
   att.global.attribute.xmllang,
   att.global.attribute.xmlbase,
   att.global.attribute.xmlspace,
   att.global.analytic.attribute.ana,
   att.global.change.attribute.change,
   att.global.facs.attribute.facs,
   att.global.linking.attribute.corresp,
   att.global.linking.attribute.synch,
```

```
   att.global.linking.attribute.sameAs,
   att.global.linking.attribute.copyOf,
   att.global.linking.attribute.next,
   att.global.linking.attribute.prev,
   att.global.linking.attribute.exclude,
   att.global.linking.attribute.select,
   att.global.rendition.attribute.rend,
   att.global.rendition.attribute.style,
   att.global.rendition.attribute.rendition,
   att.global.responsibility.attribute.cert,
   att.global.source.attribute.source,
   attribute resp { list { + } }?,
   attribute dim { "horizontal" | "vertical" }?,
   ( model.descLike | model.certLike )*
}
```

---

**** associates an interpretative annotation directly with a span of text. [*18.3. Spans and Interpretations*]

*Module* analysis — Simple Analytic Mechanisms

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*

  - * *@resp*
  - att.global.source
  - * *@source*
- att.cmc
  - *@generatedBy*
- att.interpLike
  - *@type*
  - *@subtype*
  - *@inst*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*

`@from` gives the identifier of the node which is the starting point of the span of text being annotated; if not accompanied by a *to* attribute, gives the identifier of the node of the entire span of text being annotated.
*Status* Optional
*Datatype* teidata.pointer

`@to` gives the identifier of the node which is the end-point of the span of text being annotated.
*Status* Optional
*Datatype* teidata.pointer

*Member of* model.global.meta

*Contained by*

*analysis:* cl m phr s span spanGrp w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data

*Example*

```
<p xml:id="para2">(The "aftermath" starts here)</p>
<p xml:id="para3">(The "aftermath" continues here)</p>
<p xml:id="para4">(The "aftermath" ends in this paragraph)</p>
<!-- ... -->
<span type="structure" from="#para2"
  to="#para4">aftermath</span>
```

*Schematron* <sch:rule context="tei:span"> <sch:report test="@from and @target"> Only one of the attributes @target and @from may be supplied on <sch:name/> </sch:report> </sch:rule>

*Schematron* <sch:rule context="tei:span"> <sch:report test="@to and @target"> Only one of the attributes @target and @to may be supplied on <sch:name/> </sch:report> </sch:rule>

*Schematron* <sch:rule context="tei:span"> <sch:report test="@to and not(@from)"> If @to is supplied on <sch:name/>, @from must be supplied as well </sch:report> </sch:rule>

*Schematron* <sch:rule context="tei:span"> <sch:report test="contains(normalize-space(@to),' ') or contains(normalize-space(@from),' ')"> The attributes @to and @from on <sch:name/> may each contain only a single value </sch:report> </sch:rule>

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
```

```
</content>
```

*Schema Declaration*

```
element span
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.interpLike.attributes,
   att.pointing.attributes,
   attribute from { text }?,
   attribute to { text }?,
   macro.phraseSeq.limited}
```

---

## <spanGrp> (span group) collects together span tags. [*18.3. Spans and Interpretations*]

*Module* analysis — Simple Analytic Mechanisms

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

    * *@rend*
    * *@style*
    * *@rendition*
  – att.global.responsibility
    * *@cert*
    * *@resp*
  – att.global.source
    * *@source*
- att.cmc
  – *@generatedBy*
- att.interpLike
  – *@type*
  – *@subtype*
  – *@inst*

*Member of* model.global.meta

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* annotationBlock u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*analysis:* span

*core:* desc

*Example*

```
<u xml:id="UU1">Can I have ten oranges and a kilo of bananas please?</u>
<u xml:id="UU2">Yes, anything else?</u>
<u xml:id="UU3">No thanks.</u>
<u xml:id="UU4">That'll be dollar forty.</u>
<u xml:id="UU5">Two dollars</u>
<u xml:id="UU6">Sixty, eighty, two dollars.
<anchor xml:id="UU6e"/>Thank you.<anchor xml:id="UU6f"/>
</u>
<spanGrp type="transactions">
  <span from="#UU1">sale request</span>
  <span from="#UU2" to="#UU3">sale compliance</span>
  <span from="#UU4">sale</span>
  <span from="#UU5" to="#UU6">purchase</span>
  <span from="#UU6e" to="#UU6f">purchase closure</span>
</spanGrp>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.descLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="span" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element spanGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.interpLike.attributes,
   ( model.descLike*, span* )
}
```

---

**<speaker>** contains a specialized form of heading or label, giving the name of one or more speakers in a dramatic text or fragment. [*3.13.2. Core Tags for Drama*]

*Module* core — Elements Available in All TEI Documents
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  * *@ana*
- att.global.change
  * *@change*
- att.global.facs
  * *@facs*
- att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

• att.placement
- *@place*

• att.written
- *@hand*

*Contained by* core: sp
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* This element may be used to transcribe which character is speaking in a dramatic text as indicated by the source text; the *who* attribute of an <sp> element may be used to point to another element (typically a <role>) which provides information about the character speaking. Either or both may be used.

*Example*

```
<sp  who="#ni #rsa">
  <speaker>Nancy and Robert</speaker>
  <stage  type="delivery">(speaking simultaneously)</stage>
  <p>The future? ...</p>
</sp>
<list  type="speakers">
  <item  xml:id="ni"/>
  <item  xml:id="rsa"/>
</list>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element speaker
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.placement.attributes,
```

```
   att.written.attributes,
   macro.phraseSeq}
```

---

**<specDesc>** (specification description) indicates that a description of the specified element, class, or macro should be included at this point within a document. [*23.1.2. Element and Attribute Descriptions*]

*Module* tagdocs — Documentation Elements

*Attributes* • att.global

- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  - * *@ana*
- – att.global.change
  - * *@change*
- – att.global.facs
  - * *@facs*
- – att.global.linking
  - * *@corresp*
  - * *@synch*
  - * *@sameAs*
  - * *@copyOf*
  - * *@next*
  - * *@prev*
  - * *@exclude*
  - * *@select*
- – att.global.rendition
  - * *@rend*
  - * *@style*
  - * *@rendition*
- – att.global.responsibility
  - * *@cert*
  - * *@resp*
- – att.global.source
  - * *@source*

`@key` (identifier) supplies the identifier of the documentary element or class for which a description is to be obtained.

*Status* Required

*Datatype* teidata.name

```
<specDesc key="emph"/>
```

*Note* The value should match the *ident* of a documentary element (that is, a member of att.identified).

`@atts` (attributes) supplies attribute names for which descriptions should additionally be obtained.

*Status* Recommended

*Datatype* 0–∞ occurrences of teidata.name separated by whitespace

```
<specDesc key="foreign"
  atts="usage xml:lang"/>
```

*Note* The attribute names listed may include both attributes inherited from a class and those defined explicitly for the associated element.

*Member of* model.specDescLike

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg specList

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain* Empty element

*Note* The description is usually displayed as a label and an item. The list of attributes may include some which are inherited by virtue of an element's class membership; descriptions for such attributes may also be retrieved using another <specDesc>, this time pointing at the relevant class.

*Example*

```
<specDesc key="orth"/>
```

*Example*

```
<specDesc key="emph"/>
```

*Example*

```
<specDesc  key="foreign"
  atts="usage xml:lang"/>
```

*Content model* <content> <empty/></content>

*Schema Declaration*

```
element specDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute key { text },
   attribute atts { list { * } }?,
   empty
}
```

---

**<specGrp>** (specification group) contains any convenient grouping of specifications for use within the current module. [*23.2. Modules and Schemas*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.oddDecl

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*tagdocs:* schemaSpec specGrp

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*core:* l lg p sp

*drama:* spGrp

*linking:* ab

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock u

*tagdocs:* classRef classSpec constraintSpec dataRef dataSpec elementRef elementSpec listRef macroRef macroSpec moduleRef moduleSpec outputRendition specGrp specGrpRef

*Note* A specification group is referenced by means of its *xml:id* attribute. The declarations it contains may be included in a <schemaSpec> or <moduleSpec> element only by reference (using a <specGrpRef> element): it may not be nested within a <moduleSpec> element.

Different ODD processors may generate representations of the specifications contained by a <specGrp> in different concrete syntaxes.

*Example*

```
<specGrp xml:id="xDAILC">
  <elementSpec ident="s">
<!-- ... -->
  </elementSpec>
  <elementSpec ident="cl">
<!-- ... -->
  </elementSpec>
  <elementSpec ident="w">
<!-- ... -->
  </elementSpec>
  <elementSpec ident="m">
<!-- ... -->
  </elementSpec>
  <elementSpec ident="c">
<!-- ... -->
  </elementSpec>
</specGrp>
```

This specification group with identifier xDAILC contains specifications for the elements <s>,<cl>,<w>, etc.

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.oddDecl"/>
  <classRef key="model.oddRef"/>
  <classRef key="model.divPart"/>
  <elementRef key="listRef"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element specGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.oddDecl | model.oddRef | model.divPart | listRef )*
}
```

---

**<specGrpRef>** (reference to a specification group) indicates that the declarations contained by the <specGrp> referenced should be inserted at this point. [*23.2. Modules and Schemas*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@target` points at the specification group which logically belongs here.

*Status* Required

*Datatype* teidata.pointer

*Member of* model.oddDecl

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*tagdocs:* schemaSpec specGrp

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain* Empty element

*Note* In ODD documentation processing, a <specGrpRef> usually produces a comment indicating that a set of declarations printed in another section will be inserted at this point in the <specGrp> being discussed. In schema processing, the contents of the specified <specGrp> are made available for inclusion in the generated schema.

The specification group identified by the *target* attribute will normally be part of the current ODD document.

*Example*

```
<p>This part of the module contains declarations for names of persons, places,
and
organisations: <specGrpRef  target="#names.pers"/>
  <specGrpRef  target="#names.place"/>
  <specGrpRef  target="#names.org"/>
</p>
<!-- elsewhere -->
<specGrp  xml:id="names.pers">
<!--... -->
</specGrp>
<!-- elsewhere -->
<specGrp  xml:id="names.place">
<!--... -->
</specGrp>
<!-- elsewhere -->
<specGrp  xml:id="names.org">
<!--... -->
</specGrp>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element specGrpRef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute target { text },
   empty
}
```

---

**<specList>** (specification list) marks where a list of descriptions is to be inserted into the prose documentation. [*23.1.2. Element and Attribute Descriptions*]

*Module* tagdocs — Documentation Elements

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.specDescLike

*Contained by*

*analysis:* cl phr s

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*tagdocs:* specDesc

*Example*

```
<specList>
  <specDesc  key="refState"
    atts="delim length"/>
  <specDesc  key="div"/>
</specList>
```

*Content model*

```
<content>
 <elementRef key="specDesc" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element specList
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   specDesc+
}
```

**<spGrp>** (speech group) contains a group of speeches or songs in a performance text presented in a source as constituting a single unit or 'number'. [*7.2.3. Grouped Speeches*]

*Module* drama — Performance Texts

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.ascribed.directed
  - *@toWhom*
  - att.ascribed
    - *@who*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.divPart

*Contained by*

*core:* item note q quote said stage textLang
*drama:* castList epilogue performance prologue set view
*figures:* cell figure
*header:* change handNote licence scriptNote
*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote
*namesdates:* occupation
*tagdocs:* specGrp
*textcrit:* lem rdg
*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript
*transcr:* metamark
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* cb ellipsis gap gb head index lb milestone note noteGrp pb sp stage
*drama:* camera caption move sound tech view
*figures:* figure notatedMusic
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*spoken:* incident kinesic pause shift vocal writing
*textcrit:* app witDetail
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
*Example*

```
<sp>
  <speaker>FRAULEIN SCHNEIDER:</speaker>
  <p> Herr Schultz! Can I believe what I see? <stage>(HERR SCHULTZ nods
        proudly)</stage> But this is — too much to accept. So rare — so costly
—
     so luxurious. </p>
</sp>
<stage>(She sings)</stage>
<spGrp  n="4">
  <sp>
    <l>If you bought me diamonds, If you bought me pearls,</l>
    <l>If you bought me roses like some other gents</l>
    <l>Might bring to other girls,</l>
    <l>It couldn't please me more</l>
    <l>Than the gift I see -</l>
    <stage>(She takes a large pineapple out of the bag)</stage>
    <l>A pineapple for me!</l>
  </sp>
  <sp>
    <speaker>SCHULTZ:</speaker>
    <stage>(Singing) </stage>
    <l>If, in your emotion, </l>
    <l>You began to sway, </l>
    <l>Went to get some air, </l>
    <l>Or grabbed a chair </l>
    <l>To keep from fainting dead away, </l>
    <l>It couldn't please me more </l>
    <l>Than to see you cling </l>
```

```
      <l>To the pineapple I bring.</l>
    </sp>
    <sp>
      <speaker>BOTH:</speaker>
      <l>Ah, ah, ah, ah, ah, ah, ah, ah</l>
    </sp>
  <!-- ... -->
    <stage>(They dance)</stage>
  </spGrp>
  <sp>
    <speaker>FRAULEIN SCHNEIDER: </speaker>
    <p>But you must not bring me
       any more pineapples! Do you hear? It is not proper. It is a gift a
       young man would present to his lady love. It makes me blush!
    </p>
  </sp>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate minOccurs="1"
   maxOccurs="unbounded">
   <classRef key="model.global"/>
   <elementRef key="sp"/>
   <classRef key="model.stageLike"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element spGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.directed.attributes,
   att.ascribed.attributes,
   att.typed.attributes,
   ( model.headLike*, ( model.global | sp | model.stageLike )+ )
}
```

---

**<sponsor>** (sponsor) specifies the name of a sponsoring organization or institution. [*2.2.1. The Title Statement*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*

- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.canonical
  - *@key*
  - *@ref*
- att.datable
  - *@period*

- att.datable.custom
  - *@when-custom*
  - *@notBefore-custom*
  - *@notAfter-custom*
  - *@from-custom*
  - *@to-custom*
  - *@datingPoint*
  - *@datingMethod*

- att.datable.iso
  - *@when-iso*
  - *@notBefore-iso*
  - *@notAfter-iso*
  - *@from-iso*
  - *@to-iso*
- att.datable.w3c
  - *@when*
  - *@notBefore*
  - *@notAfter*
  - *@from*
  - *@to*

*Member of* model.respLike

*Contained by*

*core:* bibl monogr

*header:* editionStmt titleStmt

*msdescription:* msItem

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit

*dictionaries:* lang

*figures:* figure notatedMusic

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin character data

*Note* Sponsors give their intellectual authority to a project; they are to be distinguished from *funders* (see element <funder>), who provide the funding but do not necessarily take intellectual responsibility.

*Example*

```
<sponsor>Association for Computers and the Humanities</sponsor>
<sponsor>Association for Computational Linguistics</sponsor>
<sponsor  ref="http://www.allc.org/">Association for Literary and Linguistic
Computing</sponsor>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element sponsor
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   macro.phraseSeq.limited}
```

---

**<stage>** (stage direction) contains any kind of stage direction within a dramatic text or fragment. [*3.13.2. Core Tags for Drama 3.13. Passages of Verse or Drama 7.2.4. Stage Directions*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*

  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

- att.ascribed.directed
  - *@toWhom*

- att.ascribed
  * *@who*

- att.cmc
  - *@generatedBy*
- att.placement
  - *@place*
- att.written
  - *@hand*

`@type` indicates the kind of stage direction.

*Status* Recommended

*Datatype* 0–∞ occurrences of teidata.enumerated separated by whitespace

*Suggested values include:* **setting** describes a setting.

**entrance** describes an entrance.

**exit** describes an exit.

**business** describes stage business.

**novelistic** is a narrative, motivating stage direction.

**delivery** describes how a character speaks.

**modifier** gives some detail about a character.

**location** describes a location.

**mixed** more than one of the above

*Note* If the value mixed is used, it must be the only value. Multiple values may however be supplied if a single stage direction performs multiple functions, for example is both an entrance and a modifier.

*Member of* model.stageLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l lg meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound spGrp tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The *who* attribute may be used to indicate more precisely the person or persons participating in the action described by the stage direction.

*Example*

```
<stage  type="setting">A curtain being drawn.</stage>
<stage  type="setting">Music</stage>
<stage  type="entrance">Enter Husband as being thrown off his horse and
falls.</stage>
<!-- Middleton : Yorkshire Tragedy -->
<stage  type="exit">Exit pursued by a bear.</stage>
<stage  type="business">He quickly takes the stone out.</stage>
<stage  type="delivery">To Lussurioso.</stage>
<stage  type="novelistic">Having had enough, and embarrassed for the
family.</stage>
<!-- Lorraine Hansbury : a raisin in in the sun -->
<stage  type="modifier">Disguised as Ansaldo.</stage>
<stage  type="entrance modifier">Enter Latrocinio disguised as an
empiric</stage>
<!-- Middleton: The Widow -->
<stage  type="location">At a window.</stage>
<stage  rend="inline"  type="delivery">Aside.</stage>
```

*Example*

```
<l>Behold. <stage  n="*"  place="margin">Here the vp<lb/>per part of the
<hi>Scene</hi> open'd; when
    straight appear'd a Heauen, and all the <hi>Pure Artes</hi> sitting on
    two semi<lb/>circular ben<lb/>ches, one a<lb/>boue another: who sate thus
till the rest of the
  <hi>Prologue</hi> was spoken, which being ended, they descended in
    order within the <hi>Scene,</hi> whiles the Musicke plaid</stage> Our
Poet knowing our free hearts</l>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element stage
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.directed.attributes,
   att.ascribed.attributes,
   att.cmc.attributes,
   att.placement.attributes,
   att.written.attributes,
   attribute type
   {
      list
      {
         (
```

```
            "setting"
          | "entrance"
          | "exit"
          | "business"
          | "novelistic"
          | "delivery"
          | "modifier"
          | "location"
          | "mixed"
        )*
    }
  }?,
  macro.specialPara}
```

---

**<stamp>** (stamp) contains a word or phrase describing a stamp or similar device. [*11.3.3. Watermarks and Stamps*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*

  - att.global.source
    * *@source*
- att.datable
  - *@period*
  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.pPart.msdesc

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<rubric>Apologyticu TTVLLIANI AC IGNORATIA IN XPO IHV<lb/>
SI NON LICET<lb/>
NOBIS RO<lb/>
manii imperii <stamp>Bodleian stamp</stamp>
  <lb/>
</rubric>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element stamp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<standOff>** Functions as a container element for linked data, contextual information, and stand-off annotations embedded in a TEI document. [*17.10. The standOff Container*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*

    * *@exclude*
    * *@select*
  – att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  – att.global.responsibility
    * *@cert*
    * *@resp*
  – att.global.source
    * *@source*
- att.declaring
  – *@decls*
- att.typed
  – *@type*
  – *@subtype*

*Member of* model.resource
*Contained by*
*core:* teiCorpus
*textstructure:* TEI
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* bibl biblStruct index list listBibl note
*dictionaries:* entry entryFree superEntry
*drama:* castList
*figures:* table
*header:* biblFull listChange xenoData
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp annotation join joinGrp link linkGrp listAnnotation seg timeline
*msdescription:* msDesc
*namesdates:* listEvent listNym listObject listOrg listPerson listPlace listRelation
*nets:* eTree forest graph listForest tree
*spoken:* annotationBlock u
*textcrit:* listApp listWit
*transcr:* listTranspose substJoin zone
*Example* This example shows an encoding of morphosyntactic features similar to the encoding system used by ISO 24611 (MAF).

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!-- ... -->
  </teiHeader>
  <text>
    <body>
<!-- ... -->
      <p>
```

```
        <w  xml:id="w51">I</w>
        <w  xml:id="w52">wanna</w>
        <w  xml:id="w53">put</w>
        <w  xml:id="w54">up</w>
        <w  xml:id="w55">new</w>
        <w  xml:id="w56">wallpaper</w>
        <pc>.</pc>
      </p>
<!-- ... -->
    </body>
  </text>
  <standOff  type="morphosyntax">
    <spanGrp  type="wordForm">
      <span  target="#w51"  ana="#fs01"/>
      <span  target="#w52"  ana="#fs02"/>
      <span  target="#w52"  ana="#fs03"/>
      <span  target="#w53 #w54"  ana="#fs04"/>
      <span  target="#w55"  ana="#fs05"/>
      <span  target="#w56"  ana="#fs06"/>
    </spanGrp>
    <fs  xml:id="fs01">
      <f  name="lemma">
        <string>I</string>
      </f>
      <f  name="pos">
        <symbol  value="PP"/>
      </f>
    </fs>
    <fs  xml:id="fs02">
      <f  name="lemma">
        <string>want</string>
      </f>
      <f  name="pos">
        <symbol  value="VBP"/>
      </f>
    </fs>
    <fs  xml:id="fs03">
      <f  name="lemma">
        <string>to</string>
      </f>
      <f  name="pos">
        <symbol  value="TO"/>
      </f>
    </fs>
    <fs  xml:id="fs04">
      <f  name="lemma">
        <string>put up</string>
      </f>
      <f  name="pos">
        <symbol  value="VB"/>
      </f>
    </fs>
    <fs  xml:id="fs05">
      <f  name="lemma">
        <string>new</string>
      </f>
      <f  name="pos">
```

```
        <symbol value="JJ"/>
      </f>
    </fs>
    <fs xml:id="fs06">
      <f name="lemma">
        <string>wallpaper</string>
      </f>
      <f name="pos">
        <symbol value="NN"/>
      </f>
    </fs>
  </standOff>
</TEI>
```

*Example* This example shows an encoding of contextual information which is referred to from the main text.

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!-- ... -->
  </teiHeader>
  <standOff>
    <listPlace>
      <place xml:id="LATL">
        <placeName>Atlanta</placeName>
        <location>
          <region key="US-GA">Georgia</region>
          <country key="USA">United States of America</country>
          <geo>33.755 -84.39</geo>
        </location>
        <population when="1963"
          type="interpolatedCensus" quantity="489359"
          source="https://www.biggestuscities.com/city/atlanta-georgia"/>
      </place>
      <place xml:id="LBHM">
        <placeName>Birmingham</placeName>
        <location>
          <region key="US-AL">Alabama</region>
          <country key="USA">United States of America</country>
          <geo>33.653333 -86.808889</geo>
        </location>
        <population when="1963"
          type="interpolatedCensus" quantity="332891"
          source="https://www.biggestuscities.com/city/birmingham-alabama"/>
      </place>
    </listPlace>
  </standOff>
  <text>
    <body>
<!-- ... -->
      <p>Moreover, I am <choice>
         <sic>congnizant</sic>
         <corr>cognizant</corr>
       </choice> of the interrelatedness of all communities and
      <lb/>states. I cannot sit idly by in
<placeName ref="#LATL">Atlanta</placeName> and not be concerned about what
happens
      <lb/>in <placeName ref="#LBHM">Birmingham</placeName>.
```

```
<seg  xml:id="FQ17">Injustice anywhere is a threat to justice everywhere.</seg>
We
     <lb/>are caught in an inescapable network of mutuality, tied in a single
garment
     <lb/>of destiny. Whatever affects one directly affects all indirectly.
Never
     <lb/>again can we afford to live with the narrow, provincial
<soCalled  rendition="#Rqms">outside agitator</soCalled>
       <lb/>idea. Anyone who lives inside the United States can never be
considered
     <lb/>an outsider anywhere in this country.</p>
<!-- ... -->
   </body>
  </text>
</TEI>
```

*Schematron* <sch:rule context="tei:standOff"> <sch:assert test="@type or not(ancestor::tei:standOff)">This <sch:name/> element must have a @type attribute, since it is nested inside a <sch:name/> </sch:assert> </sch:rule>

*Content model*

```
<content>
 <classRef key="model.standOffPart"
  minOccurs="1" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element standOff
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.typed.attributes,
   model.standOffPart+
}
```

---

**<state>** (state) contains a description of some status or quality attributed to a person, place, or organization often at some specific time or for a specific date range. [*14.3.1. Basic Principles 14.3.2.1. Personal Characteristics*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

- att.datable.custom
  - *@when-custom*
  - *@notBefore-custom*
  - *@notAfter-custom*
  - *@from-custom*
  - *@to-custom*
  - *@datingPoint*
  - *@datingMethod*
- att.datable.iso
  - *@when-iso*
  - *@notBefore-iso*
  - *@notAfter-iso*

    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*

  - att.canonical
    * *@key*
    * *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.orgStateLike model.persStateLike model.placeStateLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona place placeName region residence roleName settlement sex socecStatus state surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*certainty:* precision

*core:* bibl biblStruct desc head label listBibl note noteGrp p

*header:* biblFull

*linking:* ab

*msdescription:* msDesc

*namesdates:* state

*Note* Where there is confusion between <trait> and <state> the more general purpose element <state> should be used even for unchanging characteristics. If you wish to distinguish between characteristics that are generally perceived to be time-bound states and those assumed to be fixed traits, then <trait> is available for the more static of these. The <state> element encodes characteristics which are sometimes assumed to change, often at specific times or over a date range, whereas the <trait> elements are used to record characteristics, such as eye-colour, which are less subject to change. Traits are typically, but not necessarily, independent of the volition or action of the holder.

*Example*

```
<state ref="#SCHOL" type="status">
  <label>scholar</label>
</state>
```

*Example*

```
<org>
  <orgName notAfter="1960">The Silver Beetles</orgName>
  <orgName notBefore="1960">The Beatles</orgName>
  <state type="membership" from="1960-08"
    to="1962-05">
    <desc>
```

```
        <persName>John Lennon</persName>
        <persName>Paul McCartney</persName>
        <persName>George Harrison</persName>
        <persName>Stuart Sutcliffe</persName>
        <persName>Pete Best</persName>
      </desc>
    </state>
    <state  type="membership"  notBefore="1963">
      <desc>
        <persName>John Lennon</persName>
        <persName>Paul McCartney</persName>
        <persName>George Harrison</persName>
        <persName>Ringo Starr</persName>
      </desc>
    </state>
  </org>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="precision" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate>
   <elementRef key="state" minOccurs="1"
    maxOccurs="unbounded"/>
   <sequence>
    <classRef key="model.headLike"
     minOccurs="0" maxOccurs="unbounded"/>
    <classRef key="model.pLike"
     minOccurs="1" maxOccurs="unbounded"/>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <classRef key="model.noteLike"/>
     <classRef key="model.biblLike"/>
    </alternate>
   </sequence>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.labelLike"/>
    <classRef key="model.noteLike"/>
    <classRef key="model.biblLike"/>
   </alternate>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element state
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
```

```
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   (
      precision*,
      (
         state+
       | (
            model.headLike*,
            model.pLike+,
            ( model.noteLike | model.biblLike )*
         )
       | ( model.labelLike | model.noteLike | model.biblLike )*
      )
   )
}
```

---

**<stdVals>** (standard values) specifies the format used when standardized date or number values are supplied.
[*2.3.3. The Editorial Practices Declaration 16.3.2. Declarable Elements*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.declarable
  - *@default*

*Member of* model.editorialDeclPart
*Contained by*
*header:* editorialDecl
*May contain*
*core:* p
*linking:* ab
*Example*

```
<stdVals>
  <p>All integer numbers are left-filled with zeroes to 8 digits.</p>
</stdVals>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:stdVals"/> </sch:pattern>
*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element stdVals
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   model.pLike+
}
```

**<street>** contains a full street address including any name or number identifying a building as well as the name of the street or route on which it is located. [*3.6.2. Addresses*]

*Module* core — Elements Available in All TEI Documents

*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  - * *@ana*
- – att.global.change
  - * *@change*
- – att.global.facs
  - * *@facs*
- – att.global.linking
  - * *@corresp*
  - * *@synch*
  - * *@sameAs*
  - * *@copyOf*
  - * *@next*
  - * *@prev*
  - * *@exclude*
  - * *@select*
- – att.global.rendition
  - * *@rend*
  - * *@style*
  - * *@rendition*
- – att.global.responsibility
  - * *@cert*
  - * *@resp*
- – att.global.source
  - * *@source*

*Member of* model.addrPart

*Contained by*

*core:* address

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The order and presentation of house names and numbers and street names, etc., may vary considerably in different countries. The encoding should reflect the order which is appropriate in the country concerned.

*Example*

```
<street>via della Faggiola, 36</street>
```

*Example*

```
<street>
  <name>Duntaggin</name>, 110 Southmoor Road
</street>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element street
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

**<stress>** (stress) contains the stress pattern for a dictionary headword, if given separately. [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.notated
  - *@notation*

*Member of* model.formPart

*Contained by*

*dictionaries:* form

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media

mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Usually stress information is included within pronunciation information.

*Example*

```
<form>
  <orth>alternating current</orth>
  <stress>,....'..</stress>
</form>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element stress
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.notated.attributes,
   macro.paraContent}
```

**<string>** (string value) represents the value part of a feature-value specification which contains a string. [*19.3. Other Atomic Feature Values*]

*Module* iso-fs — Feature Structures

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*

*Member of* model.featureVal.single

*Contained by*

*iso-fs:* f fvLib if vAlt vColl vDefault vLabel vMerge vNot vRange

*May contain*

*gaiji:* g

character data

*Example*

```
<f  name="greeting">
  <string>Hello, world!</string>
</f>
```

*Content model* `<content> <macroRef key="macro.xtext"/></content>`

*Schema Declaration*

```
element string
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datcat.attributes,
   macro.xtext}
```

---

## <styleDefDecl>

<styleDefDecl> (style definition language declaration) specifies the name of the formal language in which style or renditional information is supplied elsewhere in the document. The specific version of the scheme may also be supplied. [*2.3.5. The Default Style Definition Language Declaration*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*

    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.declarable
  - *@default*
- att.styleDef
  - *@scheme*
  - *@schemeVersion*

*Member of* model.encodingDescPart
*Contained by*
*header:* encodingDesc
*May contain*
*core:* p
*linking:* ab
*Example*

```
<styleDefDecl scheme="css"
  schemeVersion="2.1"/>
<!-- ... -->
<tagsDecl>
  <rendition xml:id="boldface">font-weight: bold;</rendition>
  <rendition xml:id="italicstyle">font-style: italic;</rendition>
</tagsDecl>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:styleDefDecl"/> </sch:pattern>
*Content model*

```
<content>
 <classRef key="model.pLike" minOccurs="0"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element styleDefDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
```

```
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   att.styleDef.attributes,
   model.pLike*
}
```

---

**<subc>** (subcategorization) contains subcategorization information (transitive/intransitive, countable/non-countable, etc.) [*10.3.2. Grammatical Information*]

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.lexicographic
  - *@expand*

- *@split*
- *@value*
- *@location*
- *@mergedIn*
- *@opt*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*
- att.lexicographic.normalized
  - *@norm*
  - *@orig*

*Member of* model.entryPart model.lexicalRefinement

*Contained by*

*core:* cit

*dictionaries:* dictScrap entryFree form gramGrp

*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<entry>
  <form>
    <orth>médire</orth>
  </form>
  <gramGrp>
    <subc>t ind</subc>
  </gramGrp>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element subc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   macro.paraContent}
```

---

## <subst>

(substitution) groups one or more deletions (or surplus text) with one or more additions when the combination is to be regarded as a single intervention in the text. [*12.3.1.5. Substitutions*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs

    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.transcriptional
  - *@status*
  - *@cause*
  - *@seq*

  - att.editLike
    * *@evidence*
    * *@instant*
  - att.placement
    * *@place*
  - att.written

* *@hand*

*Member of* model.pPart.editorial

*Contained by*

*analysis:* cl pc phr s span w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*core:* add cb del gb lb milestone pb

*linking:* anchor

*transcr:* fw surplus

*Example*

```
... are all included. <del  hand="#RG">It is</del>
<subst>
  <add>T</add>
  <del>t</del>
</subst>he expressed
```

*Example*

```
 that he and his Sister Miſs D — <lb/>who always lived with him, wd. be <subst>
  <del>very</del>
```

```
    <lb/>
    <add>principally</add>
  </subst> remembered in her Will.
```

*Example*

```
  <ab>τ<subst>
      <add  place="above">ῶν</add>
      <del>α</del>
    </subst>
  συνκυρόντ<subst>
      <add  place="above">ων</add>
      <del>α</del>
    </subst>
  ἐργαστηρί<subst>
      <add  place="above">ων</add>
      <del>α</del>
    </subst>
  </ab>
```

*Example*

```
  <subst>
    <del>
      <gap  reason="illegible"  quantity="5"
        unit="character"/>
    </del>
    <add>apple</add>
  </subst>
```

*Schematron* <sch:rule context="tei:subst"> <sch:assert test="child::tei:add and (child::tei:del or child::tei:surplus)"> <sch:name/> must have at least one child add and at least one child del or surplus</sch:assert> </sch:rule>

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <elementRef key="add"/>
  <elementRef key="surplus"/>
  <elementRef key="del"/>
  <classRef key="model.milestoneLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element subst
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.transcriptional.attributes,
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   ( add | surplus | del | model.milestoneLike )+
}
```

---

**<substJoin>** (substitution join) identifies a series of possibly fragmented additions, deletions, or other revisions on a manuscript that combine to make up a single intervention in the text. [*12.3.1.5. Substitutions*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    - *@atLeast*
    - *@atMost*
    - *@min*
    - *@max*
    - *@confidence*
- att.pointing
  - *@targetLang*
  - *@target*
  - *@evaluate*
- att.transcriptional
  - *@status*
  - *@cause*
  - *@seq*

  - att.editLike
    - *@evidence*
    - *@instant*
  - att.placement
    - *@place*
  - att.written
    - *@hand*

*Member of* model.global.meta

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*certainty:* certainty precision respons

*core:* desc

*Example*

```
 While <del  xml:id="r112">pondering</del> thus <add  xml:id="r113">she
mus'd</add>, her pinions fann'd
<substJoin  target="#r112 #r113"/>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.descLike"/>
  <classRef key="model.certLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element substJoin
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.pointing.attributes,
   att.transcriptional.attributes,
```

```
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   ( model.descLike | model.certLike )*
}
```

---

**<summary>** contains an overview of the available information concerning some aspect of an item or object (for example, its intellectual content, history, layout, typography etc.) as a complement or alternative to the more detailed information carried by more specific elements. [*11.6. Intellectual Content*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by:* msDescription: decoDesc handDesc history layoutDesc msContents scriptDesc sealDesc typeDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<summary>This item consists of three books with a prologue and an epilogue.
</summary>
```

*Example*

```
<typeDesc>
  <summary>Uses a mixture of Roman and Black Letter types.</summary>
  <typeNote>Antiqua typeface, showing influence of Jenson's Venetian
     fonts.</typeNote>
  <typeNote>The black letter face is a variant of Schwabacher.</typeNote>
</typeDesc>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element summary
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.specialPara}
```

---

**<superEntry>** (super entry) groups a sequence of entries within any kind of lexical resource, such as a dictionary or lexicon which function as a single unit, for example a set of homographs. [*10.1. Dictionary Body and Overall Structure*]

*Deprecated* will be removed on 2027-03-07

*Module* dictionaries — Dictionaries

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.entryLike
  - *@type*
  - att.typed
  * ~~type~~
  * @subtype
- att.sortable
  - *@sortKey*

*Member of* model.entryLike model.entryPart
*Contained by*
*core:* cit
*dictionaries:* dictScrap entryFree
*drama:* castList epilogue performance prologue set
*figures:* figure
*linking:* standOff
*namesdates:* nym
*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript
*May contain*
*dictionaries:* dictScrap entry form
*Example*

```
<superEntry>
  <form>
    <orth>abandon</orth>
    <hyph>a|ban|don</hyph>
    <pron>@"band@n</pron>
  </form>
  <entry  n="1">
    <gramGrp>
      <pos>v</pos>
      <subc>T1</subc>
    </gramGrp>
    <sense  n="1">
      <def>to leave completely and for ever ... </def>
    </sense>
    <sense  n="2"/>
  </entry>
  <entry  n="2">
    <gramGrp>
      <pos>n</pos>
      <subc>U</subc>
    </gramGrp>
    <def>the state when one's feelings and actions are uncontrolled; freedom
from
        control</def>
  </entry>
</superEntry>
```

*Content model*

```
<content>
 <alternate>
  <sequence>
   <elementRef key="form" minOccurs="0"/>
   <elementRef key="entry" minOccurs="1"
    maxOccurs="unbounded"/>
  </sequence>
  <elementRef key="dictScrap"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element superEntry
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.entryLike.attributes,
   att.sortable.attributes,
   ( ( form?, entry+ ) | dictScrap )
}
```

---

**<supplied>** (supplied) signifies text supplied by the transcriber or editor for any reason; for example because the original cannot be read due to physical damage, or because of an obvious omission by the author or scribe. [*12.3.3.1. Damage, Illegibility, and Supplied Text*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

  - * *@synch*
  - * *@sameAs*
  - * *@copyOf*
  - * *@next*
  - * *@prev*
  - * *@exclude*
  - * *@select*
  - – att.global.rendition
  - * *@rend*
  - * *@style*
  - * *@rendition*
  - – att.global.responsibility
  - * *@cert*
  - * *@resp*
  - – att.global.source
  - * *@source*
- att.dimensions
  - – *@unit*
  - – *@quantity*
  - – *@extent*
  - – *@precision*
  - – *@scope*
  - – att.ranging
  - * *@atLeast*
  - * *@atMost*
  - * *@min*
  - * *@max*
  - * *@confidence*
- att.editLike
  - – *@evidence*
  - – *@instant*

`@reason` one or more words indicating why the text has had to be supplied, e.g. *overbinding*, *faded-ink*, *lost-folio*, *omitted-in-original*.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace

*Member of* model.choicePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del distinct editor ellipsis email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The <damage>, <gap>, <del>, <unclear> and <supplied> elements may be closely allied in use. See section *12.3.3.2. Use of the gap, del, damage, unclear, and supplied Elements in Combination* for discussion of which element is appropriate for which circumstance.

*Example*

```
I am dr Sr yr
<supplied  reason="illegible"
  source="#amanuensis_copy">very humble Servt</supplied>
Sydney Smith
```

*Example*

```
<supplied  reason="omitted-in-original">Dedication</supplied> to the duke of
Bejar
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element supplied
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   attribute reason { list { + } }?,
   macro.paraContent}
```

---

## <support>

<support> (support) contains a description of the materials etc. which make up the physical support for the written part of a manuscript or other object. [*11.7.1. Object Description*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic

* *@ana*

– att.global.change
* *@change*

– att.global.facs
* *@facs*

– att.global.linking
* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Contained by:* *msdescription:* supportDesc

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<objectDesc  form="roll">
  <supportDesc>
    <support> Parchment roll with <material>silk</material> ribbons.
    </support>
  </supportDesc>
</objectDesc>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element support
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.specialPara}
```

---

## <supportDesc>

(support description) groups elements describing the physical support for the written part of a manuscript or other object. [*11.7.1. Object Description*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
- att.global.analytic

* *@ana*

– att.global.change
* *@change*

– att.global.facs
* *@facs*

– att.global.linking
* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@material` (material) a short project-defined name for the material composing the majority of the support.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **paper**
**parch** (parchment)
**mixed**

*Contained by:* *msdescription:* objectDesc
*May contain*
*core:* p
*header:* extent
*linking:* ab
*msdescription:* collation condition foliation support
*Example*

```
<supportDesc>
  <support> Parchment roll with <material>silk</material> ribbons.
  </support>
</supportDesc>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="support" minOccurs="0"/>
   <elementRef key="extent" minOccurs="0"/>
   <elementRef key="foliation"
    minOccurs="0" maxOccurs="unbounded"/>
   <elementRef key="collation"
    minOccurs="0"/>
   <elementRef key="condition"
    minOccurs="0"/>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element supportDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute material { "paper" | "parch" | "mixed" }?,
   ( model.pLike+ | ( support?, extent?, foliation*, collation?, condition? ) )
}
```

---

**<surface>** defines a written surface as a two-dimensional coordinate space, optionally grouping one or more graphic representations of that space, zones of interest within that space, and, when using an embedded transcription approach, transcriptions of the writing within them. [*12.1. Digital Facsimiles 12.2.2. Embedded Transcription*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*

- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.coordinated
  - *@start*
  - *@ulx*
  - *@uly*
  - *@lrx*
  - *@lry*
  - *@points*
- att.declaring
  - *@decls*
- att.typed
  - *@type*
  - *@subtype*

`@attachment` describes the method by which this surface is or was connected to the main surface.

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **glued** glued in place

**pinned** pinned or stapled in place

**sewn** sewn in place

`@flipping` indicates whether the surface is attached and folded in such a way as to provide two writing surfaces.

*Status* Optional

*Datatype* teidata.truthValue

*Contained by:* facsimile sourceDoc surface surfaceGrp zone

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* binaryObject cb desc ellipsis gap gb graphic index label lb media milestone note noteGrp pb

*figures:* figure formula notatedMusic

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*transcr:* addSpan damageSpan delSpan fw line listTranspose metamark path space substJoin surface surfaceGrp zone

*Note* The <surface> element represents any two-dimensional space on some physical surface forming part of the source material, such as a piece of paper, a face of a monument, a billboard, a scroll, a leaf etc.

The coordinate space defined by this element may be thought of as a grid *lrx* - *ulx* units wide and *uly* - *lry* units high.

The <surface> element may contain graphic representations or transcriptions of written zones, or both. The coordinate values used by every <zone> element contained by this element are to be understood with reference to the same grid.

Where it is useful or meaningful to do so, any grouping of multiple <surface> elements may be indicated using the <surfaceGrp> element.

*Example*

```
<facsimile>
  <surface ulx="0" uly="0" lrx="200" lry="300">
    <graphic url="Bovelles-49r.png"/>
  </surface>
</facsimile>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.global"/>
   <classRef key="model.labelLike"/>
   <classRef key="model.graphicLike"/>
  </alternate>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <alternate>
    <elementRef key="zone"/>
    <elementRef key="line"/>
    <elementRef key="path"/>
    <elementRef key="surface"/>
    <elementRef key="surfaceGrp"/>
   </alternate>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element surface
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.coordinated.attributes,
   att.declaring.attributes,
   att.typed.attributes,
   attribute attachment { text }?,
   attribute flipping { text }?,
   (
      ( model.global | model.labelLike | model.graphicLike )*,
      ( ( ( zone | line | path | surface | surfaceGrp ), model.global* )* )
   )
}
```

---

**<surfaceGrp>** (surface group) defines any kind of useful grouping of written surfaces, for example the recto and verso of a single leaf, which the encoder wishes to treat as a single unit. [*12.1. Digital Facsimiles*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declaring
  - *@decls*
- att.typed
  - *@type*
  - *@subtype*

*Contained by* *transcr:* facsimile sourceDoc surface surfaceGrp
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* cb ellipsis gap gb index lb milestone note noteGrp pb
*figures:* figure notatedMusic
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*spoken:* incident kinesic pause shift vocal writing
*textcrit:* app witDetail
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin surface surfaceGrp
*Note* Where it is useful or meaningful to do so, any grouping of multiple <surface> elements may be indicated using the <surfaceGrp> elements.
*Example*

```
<sourceDoc>
  <surfaceGrp>
    <surface ulx="0" uly="0" lrx="200"
      lry="300">
      <graphic url="Bovelles-49r.png"/>
    </surface>
    <surface ulx="0" uly="0" lrx="200"
      lry="300">
      <graphic url="Bovelles-49v.png"/>
    </surface>
  </surfaceGrp>
</sourceDoc>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.global"/>
  <elementRef key="surface"/>
  <elementRef key="surfaceGrp"/>
```

```
 </alternate>
</content>
```

*Schema Declaration*

```
element surfaceGrp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.typed.attributes,
   ( model.global | surface | surfaceGrp )+
}
```

---

**<surname>** (surname) contains a family (inherited) name, as opposed to a given, baptismal, or nick name.
[*14.2.1. Personal Names*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition

    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.personal
  - *@full*
  - *@sort*
  - att.naming
    * *@role*
    * *@nymRef*
    * att.canonical
      · *@key*
      · *@ref*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.persNamePart

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit

forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<surname type="combine">St John Stevas</surname>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element surname
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.personal.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<surplus>** (surplus) marks text present in the source which the editor believes to be superfluous or redundant.
[*12.3.3.1. Damage, Illegibility, and Supplied Text*]

*Module* transcr — Representation of Primary Sources

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*

  - * *@resp*
  - att.global.source
  - * *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*
  - att.ranging
  - * *@atLeast*
  - * *@atMost*
  - * *@min*
  - * *@max*
  - * *@confidence*
- att.editLike
  - *@evidence*
  - *@instant*

`@reason` one or more words indicating why this text is believed to be superfluous, e.g. *repeated*, *interpolated* etc.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.word separated by whitespace

*Member of* model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw metamark mod restore retrace secl subst supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
I am dr Sr yrs
<surplus  reason="repeated">yrs</surplus>
Sydney Smith
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element surplus
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.editLike.attributes,
   attribute reason { list { + } }?,
   macro.paraContent}
```

---

**<surrogates>** (surrogates) contains information about any representations of the manuscript or other object being described which may exist in the holding institution or elsewhere. [*11.9. Additional Information*]

*Module* msdescription — Manuscript Description

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*

– att.global.source
* *@source*

*Contained by:* *msdescription:* additional
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound spGrp tech view
*figures:* figure formula notatedMusic table
*gaiji:* g
*header:* biblFull idno
*iso-fs:* fLib fs fvLib
*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*nets:* eTree forest graph listForest tree
*spoken:* annotationBlock incident kinesic pause shift u vocal writing
*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val
*textcrit:* app listApp listWit witDetail
*textstructure:* floatingText
*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo
*verse:* caesura rhyme
character data
*Example*

```
<surrogates>
  <bibl>
    <title  type="gmd">diapositive</title>
    <idno>AM 74 a, fol.</idno>
    <date>May 1984</date>
  </bibl>
  <bibl>
    <title  type="gmd">b/w prints</title>
    <idno>AM 75 a, fol.</idno>
    <date>1972</date>
  </bibl>
</surrogates>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element surrogates
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.specialPara}
```

---

**<syll>** (syllabification) contains the syllabification of the headword. [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.lexicographic.normalized
    * *@norm*
    * *@orig*
- att.notated
  - *@notation*

*Member of* model.entryPart model.formPart

*Contained by*

*core:* cit

*dictionaries:* dictScrap entryFree form

*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<form>
  <orth>area</orth>
  <hyph>ar|ea</hyph>
  <syll>ar|e|a</syll>
</form>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element syll
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.notated.attributes,
   macro.paraContent}
```

---

## <symbol>

<symbol> (symbolic value) represents the value part of a feature-value specification which contains one of a finite list of symbols. [*19.3. Other Atomic Feature Values*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*

`@value` supplies a symbolic value for the feature, one of a finite list that may be specified in a feature declaration.
*Status* Required
*Datatype* teidata.word

*Member of* model.featureVal.single
*Contained by*
*iso-fs:* f fvLib if vAlt vColl vDefault vLabel vMerge vNot vRange
*May contain* Empty element
*Example*

```
<f name="gender">
```

```
      <symbol value="feminine"/>
    </f>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element symbol
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datcat.attributes,
   attribute value { text },
   empty
}
```

---

## <table> (table) contains text displayed in tabular form, in rows and columns. [*15.1.1. TEI Tables*]

*Module* figures — Tables, Formulæ, Graphics, and Notated Music

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

  - * *@rend*
  - * *@style*
  - * *@rendition*
  - – att.global.responsibility
  - * *@cert*
  - * *@resp*
  - – att.global.source
  - * *@source*
- att.cmc
  - – *@generatedBy*
- att.typed
  - – *@type*
  - – *@subtype*

`@rows` (rows) indicates the number of rows in the table.
*Status* Optional
*Datatype* teidata.count
*Note* If no number is supplied, an application must calculate the number of rows.
Rows should be presented from top to bottom.

`@cols` (columns) indicates the number of columns in each row of the table.
*Status* Optional
*Datatype* teidata.count
*Note* If no number is supplied, an application must calculate the number of columns.
Within each row, columns should be presented left to right.

*Member of* model.listLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound tech view

*figures:* cell figDesc figure

*header:* abstract change handNote licence rendition scriptNote sourceDesc tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument back body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons
*core:* binaryObject cb ellipsis gap gb graphic head index lb media meeting milestone note noteGrp pb
*figures:* figure formula notatedMusic row
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*spoken:* incident kinesic pause shift vocal writing
*textcrit:* app witDetail
*textstructure:* argument byline closer dateline docAuthor docDate epigraph postscript salute signed trailer
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
*Note* Contains an optional heading and a series of rows.

Any rendition information should be supplied using the global *rend* attribute, at the table, row, or cell level as appropriate.

*Example*

```
<table rows="4" cols="4">
  <head>Poor Men's Lodgings in Norfolk (Mayhew, 1843)</head>
  <row role="label">
    <cell role="data"/>
    <cell role="data">Dossing Cribs or Lodging Houses</cell>
    <cell role="data">Beds</cell>
    <cell role="data">Needys or Nightly Lodgers</cell>
  </row>
  <row role="data">
    <cell role="label">Bury St Edmund's</cell>
    <cell role="data">5</cell>
    <cell role="data">8</cell>
    <cell role="data">128</cell>
  </row>
  <row role="data">
    <cell role="label">Thetford</cell>
    <cell role="data">3</cell>
    <cell role="data">6</cell>
    <cell role="data">36</cell>
  </row>
  <row role="data">
    <cell role="label">Attleboro'</cell>
    <cell role="data">3</cell>
    <cell role="data">5</cell>
    <cell role="data">20</cell>
  </row>
  <row role="data">
    <cell role="label">Wymondham</cell>
    <cell role="data">1</cell>
    <cell role="data">11</cell>
    <cell role="data">22</cell>
  </row>
</table>
```

*Content model*

```
<content>
 <sequence>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.headLike"/>
   <classRef key="model.global"/>
```

```
  </alternate>
  <alternate>
   <sequence minOccurs="1"
    maxOccurs="unbounded">
    <elementRef key="row"/>
    <classRef key="model.global"
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
   <sequence minOccurs="1"
    maxOccurs="unbounded">
    <classRef key="model.graphicLike"/>
    <classRef key="model.global"
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
  </alternate>
  <sequence minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.divBottom"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element table
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.typed.attributes,
   attribute rows { text }?,
   attribute cols { text }?,
   (
      ( model.headLike | model.global )*,
      (
         ( ( row, model.global* )+ )
       | ( ( model.graphicLike, model.global* )+ )
      ),
      ( ( model.divBottom, model.global* )* )
   )
}
```

---

**<tag>** (tag) contains text of a complete start- or end-tag, possibly including attribute specifications, but excluding the opening and closing markup delimiter characters. [*23. Documentation Elements*]

*Module* tagdocs — Documentation Elements
*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.typed
  - ~~type~~
  - @subtype

`@type` indicates the type of XML tag intended.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **start** a start-tag, with delimiters < and > is intended
**end** an end-tag, with delimiters </ and > is intended
**empty** an empty tag, with delimiters < and /> is intended
**pi** a pi (processing instruction), with delimiters <? and ?> is intended
**comment** a comment, with delimiters <!– and –> is intended
**ms** a marked-section, with delimiters <[CDATA[ and ]]> is intended

`@scheme` supplies the name of the schema in which this tag is defined.

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **TEI** (text encoding initiative) This tag is defined as part of the TEI scheme.[Default]

**DBK** (docbook) this tag is part of the Docbook scheme.

**XX** (unknown) this tag is part of an unknown scheme.

**Schematron**

**HTML**

*Member of* model.phrase.xml

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain* Character data only

*Example*

```
Mark the start of each italicised phrase with a
<tag>hi rend="it"</tag> tag, and its end with a <tag  type="end">hi</tag> tag.
<tag  type="comment">Example updated on 2008-04-05</tag>
```

*Content model* `<content> <textNode/></content>`

*Schema Declaration*

```
element tag
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { "start" | "end" | "empty" | "pi" | "comment" | "ms" }?,
   attribute scheme { text }?,
   text
}
```

---

**<tagsDecl>** (tagging declaration) provides detailed information about the tagging applied to a document. [*2.3.4. The Tagging Declaration 2.3. The Encoding Description*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*

* *@rendition*

– att.global.responsibility

* *@cert*

* *@resp*

– att.global.source

* *@source*

`@partial` indicates whether the element types listed exhaustively include all those found within <text>, or represent only a subset.

*Status* Recommended

*Datatype* teidata.truthValue

*Note* TEI recommended practice is to specify this attribute. When the <tagUsage> elements inside <tagsDecl> are used to list each of the element types in the associated <text>, the value should be given as false. When the <tagUsage> elements inside <tagsDecl> are used to provide usage information or default renditions for only a subset of the elements types within the associated <text>, the value should be true.

*Member of* model.encodingDescPart

*Contained by*

*header:* encodingDesc

*May contain*

*header:* namespace rendition

*Example*

```
<tagsDecl partial="true">
  <rendition xml:id="rend-it" scheme="css"
    selector="emph, hi, name, title">font-style: italic;</rendition>
  <namespace name="http://www.tei-c.org/ns/1.0">
    <tagUsage gi="hi" occurs="467"/>
    <tagUsage gi="title" occurs="45"/>
  </namespace>
  <namespace name="http://docbook.org/ns/docbook">
    <tagUsage gi="para" occurs="10"/>
  </namespace>
</tagsDecl>
```

If the *partial* attribute were not specified here, the implication would be that the document in question contains only <hi>, <title>, and `<para>` elements.

*Content model*

```
<content>
 <sequence>
  <elementRef key="rendition" minOccurs="0"
   maxOccurs="unbounded"/>
  <elementRef key="namespace" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element tagsDecl
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute partial { text }?,
   ( rendition*, namespace* )
}
```

---

**<tagUsage>** (element usage) documents the usage of a specific element within a specified document. [*2.3.4. The Tagging Declaration*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

  - *@source*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*

`@gi` (generic identifier) specifies the name (generic identifier) of the element indicated by the tag, within the namespace indicated by the parent <namespace> element.
*Status* Required
*Datatype* teidata.name

`@occurs` specifies the number of occurrences of this element within the text.
*Status* Recommended
*Datatype* teidata.count

`@withId` (with unique identifier) specifies the number of occurrences of this element within the text which bear a distinct value for the global *xml:id* attribute.
*Status* Recommended
*Datatype* teidata.count

*Contained by:* namespace
*May contain*
*core:* abbr address bibl biblStruct choice cit date desc distinct email emph expan foreign gloss hi label list listBibl measure measureGrp mentioned name num ptr q quote ref rs said soCalled stage term time title unit
*dictionaries:* lang
*drama:* camera caption castList move sound tech view
*figures:* table
*header:* biblFull idno
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specGrp specGrpRef tag val
*textcrit:* listApp listWit
*textstructure:* floatingText
*transcr:* am ex subst
character data
*Example*

```
<tagsDecl partial="true">
  <rendition xml:id="it" scheme="css"
    selector="foreign, hi"> font-style: italic; </rendition>
<!-- ... -->
  <namespace name="http://www.tei-c.org/ns/1.0">
    <tagUsage gi="hi" occurs="28" withId="2"> Used to mark English words
italicized in the copy text.</tagUsage>
    <tagUsage gi="foreign">Used to mark non-English words in the copy
text.</tagUsage>
```

```
    <!-- ... -->
      </namespace>
    </tagsDecl>
```

*Content model*

```
<content>
 <macroRef key="macro.limitedContent"/>
</content>
```

*Schema Declaration*

```
element tagUsage
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.datcat.attributes,
   attribute gi { text },
   attribute occurs { text }?,
   attribute withId { text }?,
   macro.limitedContent}
```

---

**<taxonomy>** (taxonomy) defines a typology either implicitly, by means of a bibliographic citation, or explicitly by a structured taxonomy. [*2.3.7. The Classification Declaration*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*

*Contained by:* *header:* classDecl taxonomy
*May contain*
*core:* bibl biblStruct desc gloss listBibl
*header:* biblFull category taxonomy
*msdescription:* msDesc
*tagdocs:* equiv
*Note* Nested taxonomies are common in many fields, so the <taxonomy> element can be nested.
*Example*

```
<taxonomy xml:id="tax.b">
  <bibl>Brown Corpus</bibl>
  <category xml:id="tax.b.a">
    <catDesc>Press Reportage</catDesc>
    <category xml:id="tax.b.a1">
      <catDesc>Daily</catDesc>
    </category>
    <category xml:id="tax.b.a2">
      <catDesc>Sunday</catDesc>
    </category>
    <category xml:id="tax.b.a3">
      <catDesc>National</catDesc>
    </category>
    <category xml:id="tax.b.a4">
      <catDesc>Provincial</catDesc>
    </category>
    <category xml:id="tax.b.a5">
      <catDesc>Political</catDesc>
    </category>
    <category xml:id="tax.b.a6">
      <catDesc>Sports</catDesc>
```

```
      </category>
    </category>
    <category xml:id="tax.b.d">
      <catDesc>Religion</catDesc>
      <category xml:id="tax.b.d1">
        <catDesc>Books</catDesc>
      </category>
      <category xml:id="tax.b.d2">
        <catDesc>Periodicals and tracts</catDesc>
      </category>
    </category>
  </taxonomy>
```

*Example*

```
  <taxonomy>
    <category xml:id="literature">
      <catDesc>Literature</catDesc>
      <category xml:id="poetry">
        <catDesc>Poetry</catDesc>
        <category xml:id="sonnet">
          <catDesc>Sonnet</catDesc>
          <category xml:id="shakesSonnet">
            <catDesc>Shakespearean Sonnet</catDesc>
          </category>
          <category xml:id="petraSonnet">
            <catDesc>Petrarchan Sonnet</catDesc>
          </category>
        </category>
        <category xml:id="haiku">
          <catDesc>Haiku</catDesc>
        </category>
      </category>
      <category xml:id="drama">
        <catDesc>Drama</catDesc>
      </category>
    </category>
    <category xml:id="meter">
      <catDesc>Metrical Categories</catDesc>
      <category xml:id="feet">
        <catDesc>Metrical Feet</catDesc>
        <category xml:id="iambic">
          <catDesc>Iambic</catDesc>
        </category>
        <category xml:id="trochaic">
          <catDesc>trochaic</catDesc>
        </category>
      </category>
      <category xml:id="feetNumber">
        <catDesc>Number of feet</catDesc>
        <category xml:id="pentameter">
          <catDesc>>Pentameter</catDesc>
        </category>
        <category xml:id="tetrameter">
          <catDesc>>Tetrameter</catDesc>
        </category>
      </category>
```

```
  </category>
</taxonomy>
<!-- elsewhere in document -->
<lg  ana="#shakesSonnet #iambic #pentameter">
  <l>Shall I compare thee to a summer's day</l>
<!-- ... -->
</lg>
```

*Content model*

```
<content>
 <alternate>
  <alternate>
   <alternate minOccurs="1"
    maxOccurs="unbounded">
    <elementRef key="category"/>
    <elementRef key="taxonomy"/>
   </alternate>
   <sequence>
    <alternate minOccurs="1"
     maxOccurs="unbounded">
     <classRef key="model.descLike"
      minOccurs="1" maxOccurs="1"/>
     <elementRef key="equiv" minOccurs="1"
      maxOccurs="1"/>
     <elementRef key="gloss" minOccurs="1"
      maxOccurs="1"/>
    </alternate>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <elementRef key="category"/>
     <elementRef key="taxonomy"/>
    </alternate>
   </sequence>
  </alternate>
  <sequence>
   <classRef key="model.biblLike"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="category"/>
    <elementRef key="taxonomy"/>
   </alternate>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element taxonomy
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.datcat.attributes,
   (
      (
         ( category | taxonomy )+
       | ( ( model.descLike | equiv | gloss )+, ( category | taxonomy )* )
      )
    | ( model.biblLike, ( category | taxonomy )* )
   )
}
```

---

**<tech>** (technical stage direction) describes a special-purpose stage direction that is not meant for the actors. [*7.3.1. Technical Information*]

*Module* drama — Performance Texts

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

  * *@source*

- att.typed
  - ~~type~~
  - @subtype

`@type` categorizes the technical stage direction.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **light** a lighting cue
**sound** a sound cue
**prop** a prop cue
**block** a blocking instruction

`@perf` (performance) points to one or more <performance> elements documenting the performance or performances to which this technical direction applies.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Member of* model.stageLike
*Contained by*
*cmc:* post
*core:* add corr del desc emph head hi item l lg meeting note orig p q quote ref reg said sic sp stage textLang title unclear
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr
*drama:* camera caption castList epilogue performance prologue set sound spGrp tech view
*figures:* cell figDesc figure
*header:* change handNote licence rendition scriptNote tagUsage
*iso-fs:* fDescr fsDescr
*linking:* ab seg
*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote
*namesdates:* occupation
*spoken:* u writing
*textcrit:* lem rdg witness
*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer
*transcr:* damage metamark mod restore retrace secl supplied surplus
*verse:* rhyme
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound tech view
*figures:* figure formula notatedMusic table

*gaiji:* g
*header:* biblFull idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*spoken:* incident kinesic pause shift vocal writing
*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val
*textcrit:* app listApp listWit witDetail
*textstructure:* floatingText
*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo
*verse:* caesura rhyme

character data

*Example*

```
<tech type="light">Red spot on his face</tech>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element tech
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { "light" | "sound" | "prop" | "block" }?,
   attribute perf { list { + } }?,
   macro.paraContent}
```

---

**<TEI>** (TEI document) contains a single TEI-conformant document, combining a single TEI header with one or more members of the `model.resource` class. Multiple <TEI> elements may be combined within a <TEI> (or <teiCorpus>) element. [*4. Default Text Structure 16.1. Varieties of Composite Text*]

*Module* textstructure — Default Text Structure

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.typed
  - *@type*
  - *@subtype*

`@version` specifies the version number of the TEI Guidelines against which this document is valid.

*Status* Optional

*Datatype* teidata.version

*Note* Major editions of the Guidelines have long been informally referred to by a name made up of the letter P (for Proposal) followed by a digit. The current release is one of the many releases of the fifth major edition of the Guidelines, known as P5. This attribute may be used to associate a TEI document with a specific release of the P5 Guidelines, in the absence of a more precise association provided by the *source* attribute on the associated <schemaSpec>.

*Member of* model.describedResource

*Contained by*
*core:* teiCorpus
*textstructure:* TEI
*May contain*
*header:* teiHeader
*iso-fs:* fsdDecl
*linking:* standOff
*textstructure:* TEI text
*transcr:* facsimile sourceDoc

*Note* As with all elements in the TEI scheme (except <egXML>) this element is in the TEI namespace (see *v.7.2 Namespaces*). Thus, when it is used as the outermost element of a TEI document, it is necessary to specify the TEI namespace on it. This is customarily achieved by including `http://www.tei-c.org/ns/1.0` as the value of the XML namespace declaration (xmlns), without indicating a prefix, and then not using a prefix on TEI elements in the rest of the document. For example: <TEI version="4.8.1" xml:lang="it" xmlns="http://www.tei-c.org/ns/1.0">.

*Example*

```
<TEI  version="3.3.0" xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>The shortest TEI Document Imaginable</title>
      </titleStmt>
      <publicationStmt>
        <p>First published as part of TEI P2, this is the P5
                version using a namespace.</p>
      </publicationStmt>
      <sourceDesc>
        <p>No source: this is an original work.</p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
    <body>
      <p>This is about the shortest TEI document imaginable.</p>
    </body>
  </text>
</TEI>
```

*Example*

```
<TEI  version="2.9.1" xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>A TEI Document containing four page images </title>
      </titleStmt>
      <publicationStmt>
        <p>Unpublished demonstration file.</p>
      </publicationStmt>
      <sourceDesc>
        <p>No source: this is an original work.</p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
```

```
    <facsimile>
      <graphic  url="page1.png"/>
      <graphic  url="page2.png"/>
      <graphic  url="page3.png"/>
      <graphic  url="page4.png"/>
    </facsimile>
  </TEI>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="teiHeader"/>
  <alternate>
   <sequence>
    <classRef key="model.resource"
     minOccurs="1" maxOccurs="unbounded"/>
    <elementRef key="TEI" minOccurs="0"
     maxOccurs="unbounded"/>
   </sequence>
   <elementRef key="TEI" minOccurs="1"
    maxOccurs="unbounded"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element TEI
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   attribute version { text }?,
   ( teiHeader, ( ( model.resource+, TEI* ) | TEI+ ) )
}
```

---

## <teiCorpus>

(TEI corpus) contains the whole of a TEI encoded corpus, comprising a single corpus header and one or more <TEI> elements, each containing a single text header and a text. [*4. Default Text Structure 16.1. Varieties of Composite Text*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.typed
  - *@type*
  - *@subtype*

`@version` (version) specifies the version number of the TEI Guidelines against which this document is valid.

*Status* Optional

*Datatype* teidata.version

*Note* Major editions of the Guidelines have long been informally referred to by a name made up of the letter P (for Proposal) followed by a digit. The current release is one of the many releases of the fifth major edition of the Guidelines, known as P5. This attribute may be used to associate a TEI document with a specific release of the P5 Guidelines, in the absence of a more precise association provided by the *source* attribute on the associated <schemaSpec>.

*Member of* model.describedResource

*Contained by*

*core:* teiCorpus

*May contain*

*core:* teiCorpus

*header:* teiHeader

*iso-fs:* fsdDecl

*linking:* standOff

*textstructure:* TEI text

*transcr:* facsimile sourceDoc

*Note* Should contain one <teiHeader> for the corpus, and a series of <TEI> elements, one for each text.

As with all elements in the TEI scheme (except <egXML>) this element is in the TEI namespace (see *v.7.2 Namespaces*). Thus, when it is used as the outermost element of a TEI document, it is necessary to specify the TEI namespace on it. This is customarily achieved by including `http://www.tei-c.org/ns/1.0` as the value of the XML namespace declaration (xmlns), without indicating a prefix, and then not using a prefix on TEI elements in the rest of the document. For example: <teiCorpus version="4.8.1" xml:lang="en" xmlns="http://www.tei-c.org/ns/1.0">.

*Example*

```
<teiCorpus  version="3.3.0" xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
<!-- header for corpus -->
  </teiHeader>
  <TEI>
    <teiHeader>
<!-- header for first text -->
    </teiHeader>
    <text>
<!-- content of first text -->
    </text>
  </TEI>
  <TEI>
    <teiHeader>
<!-- header for second text -->
    </teiHeader>
    <text>
<!-- content of second text -->
    </text>
  </TEI>
<!-- more TEI elements here -->
</teiCorpus>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="teiHeader"/>
  <classRef key="model.resource"
   minOccurs="0" maxOccurs="unbounded"/>
  <classRef key="model.describedResource"
   minOccurs="1" maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element teiCorpus
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
```

```
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attributes,
   attribute version { text }?,
   ( teiHeader, model.resource*, model.describedResource+ )
}
```

---

**<teiHeader>** (TEI header) supplies descriptive and declarative metadata associated with a digital resource or set of resources. [*2.1.1. The TEI Header and Its Components 16.1. Varieties of Composite Text*]

*Module* header — The TEI Header

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by:* teiCorpus
*textstructure:* TEI
*May contain*
*header:* encodingDesc fileDesc profileDesc revisionDesc xenoData
*Note* One of the few elements unconditionally required in any TEI document.
*Example*

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>Shakespeare: the first folio (1623) in electronic form</title>
      <author>Shakespeare, William (1564–1616)</author>
      <respStmt>
        <resp>Originally prepared by</resp>
        <name>Trevor Howard-Hill</name>
      </respStmt>
      <respStmt>
        <resp>Revised and edited by</resp>
        <name>Christine Avern-Carr</name>
      </respStmt>
    </titleStmt>
    <publicationStmt>
      <distributor>Oxford Text Archive</distributor>
      <address>
        <addrLine>13 Banbury Road, Oxford OX2 6NN, UK</addrLine>
      </address>
      <idno  type="OTA">119</idno>
      <availability>
        <p>Freely available on a non-commercial basis.</p>
      </availability>
      <date  when="1968">1968</date>
    </publicationStmt>
    <sourceDesc>
      <bibl>The first folio of Shakespeare, prepared by Charlton Hinman (The
Norton Facsimile,
           1968)</bibl>
    </sourceDesc>
  </fileDesc>
  <encodingDesc>
    <projectDesc>
      <p>Originally prepared for use in the production of a series of
old-spelling
           concordances in 1968, this text was extensively checked and revised
for use during the
           editing of the new Oxford Shakespeare (Wells and Taylor, 1989).</p>
    </projectDesc>
    <editorialDecl>
      <correction>
        <p>Turned letters are silently corrected.</p>
      </correction>
      <normalization>
        <p>Original spelling and typography is retained, except that long s and
ligatured
              forms are not encoded.</p>
      </normalization>
    </editorialDecl>
    <refsDecl  xml:id="ASLREF">
```

```
      <cRefPattern  matchPattern="(\S+) ([^.]+)\.(.*)"
       replacementPattern="#xpath(//div1[@n='$1']/div2/[@n='$2']//lb[@n='$3'])">
       <p>A reference is created by assembling the following, in the reverse
order as that
            listed here: <list>
          <item>the <att>n</att> value of the preceding <gi>lb</gi>
          </item>
          <item>a period</item>
          <item>the <att>n</att> value of the ancestor <gi>div2</gi>
          </item>
          <item>a space</item>
          <item>the <att>n</att> value of the parent <gi>div1</gi>
          </item>
        </list>
       </p>
      </cRefPattern>
    </refsDecl>
  </encodingDesc>
  <revisionDesc>
    <list>
      <item>
       <date  when="1989-04-12">12 Apr 89</date> Last checked by CAC</item>
      <item>
       <date  when="1989-03-01">1 Mar 89</date> LB made new file</item>
    </list>
  </revisionDesc>
</teiHeader>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="fileDesc"/>
  <classRef key="model.teiHeaderPart"
   minOccurs="0" maxOccurs="unbounded"/>
  <elementRef key="revisionDesc"
   minOccurs="0"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element teiHeader
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( fileDesc, model.teiHeaderPart*, revisionDesc? )
}
```

**<term>** (term) contains a single-word, multi-word, or symbolic designation which is regarded as a technical term. [*3.4.1. Terms and Glosses*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cReferencing
  - *@cRef*
- att.canonical
  - *@key*
  - *@ref*
- att.cmc
  - *@generatedBy*
- att.declaring
  - *@decls*
- att.pointing

- *@targetLang*
- *@target*
- *@evaluate*

- att.sortable
  - *@sortKey*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.emphLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi index item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote keywords language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* When this element appears within an <index> element, it is understood to supply the form under which an index entry is to be made for that location. Elsewhere, it is understood simply to indicate that its content is to be regarded as a technical or specialised term. It may be associated with a <gloss> element by means of its *ref* attribute; alternatively a <gloss> element may point to a <term> element by means of its *target* attribute.

In formal terminological work, there is frequently discussion over whether terms must be atomic or may include multi-word lexical items, symbolic designations, or phraseological units. The <term> element may be used to mark any of these. No position is taken on the philosophical issue of what a term can be; the looser definition simply allows the <term> element to be used by practitioners of any persuasion.

As with other members of the `att.canonical` class, instances of this element occuring in a text may be associated with a canonical definition, either by means of a URI (using the *ref* attribute), or by means of some system-specific code value (using the *key* attribute). Because the mutually exclusive *target* and *cRef* attributes overlap with the function of the *ref* attribute, they are deprecated and may be removed at a subsequent release.

*Example*

```
A computational device that infers structure
from grammatical strings of words is known as a <term>parser</term>, and much
of the history
of NLP over the last 20 years has been occupied with the design of parsers.
```

*Example*

```
We may define <term  xml:id="TDPV1"  rend="sc">discoursal point of view</term>
as
<gloss  target="#TDPV1">the relationship, expressed
through discourse structure, between the implied author or some other
addresser, and the
fiction.</gloss>
```

*Example*

```
We may define <term  ref="#TDPV2"  rend="sc">discoursal point of view</term> as
```

```
<gloss  xml:id="TDPV2">the relationship, expressed
through discourse structure, between the implied author or some other
addresser, and the
fiction.</gloss>
```

*Example*

```
We discuss Leech's concept of
<term  ref="myGlossary.xml#TDPV2"  rend="sc">discoursal point of view</term>
below.
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element term
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cReferencing.attributes,
   att.canonical.attributes,
   att.cmc.attributes,
   att.declaring.attributes,
   att.pointing.attributes,
   att.sortable.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<terrain>** (terrain) contains information about the physical terrain of a place. [*14.3.4.3. States, Traits, and Events*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change

    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*

    - *@notAfter*
    - *@from*
    - *@to*
  - att.editLike
    - *@evidence*
    - *@instant*
  - att.naming
    - *@role*
    - *@nymRef*
  - att.canonical
    - *@key*
    - *@ref*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.placeStateLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns place placeName region residence roleName settlement sex socecStatus surname terrain

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme
*May contain*
*certainty:* precision
*core:* bibl biblStruct desc head label listBibl note noteGrp p
*header:* biblFull
*linking:* ab
*msdescription:* msDesc
*namesdates:* terrain
*Example*

```
<place  xml:id="KERG">
  <placeName>Kerguelen Islands</placeName>
<!-- ... -->
  <terrain>
    <desc>antarctic tundra</desc>
  </terrain>
<!-- ... -->
</place>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="precision" minOccurs="0"
   maxOccurs="unbounded"/>
  <classRef key="model.headLike"
   minOccurs="0" maxOccurs="unbounded"/>
  <alternate>
   <classRef key="model.pLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <classRef key="model.labelLike"
    minOccurs="1" maxOccurs="unbounded"/>
  </alternate>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.noteLike"/>
   <classRef key="model.biblLike"/>
  </alternate>
  <elementRef key="terrain" minOccurs="0"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element terrain
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
```

```
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   (
      precision*,
      model.headLike*,
      ( model.pLike+ | model.labelLike+ ),
      ( model.noteLike | model.biblLike )*,
      terrain*
   )
}
```

---

**<text>** (text) contains a single text of any kind, whether unitary or composite, for example a poem or drama, a collection of essays, a novel, a dictionary, or a corpus sample. [*4. Default Text Structure 16.1. Varieties of Composite Text*]

*Module* textstructure — Default Text Structure

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*

    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.declaring
  - *@decls*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*

*Member of* model.resource

*Contained by*

*core:* teiCorpus

*textstructure:* TEI group

*May contain*

*analysis:* interp interpGrp span spanGrp

*certainty:* certainty precision respons

*core:* cb ellipsis gap gb index lb milestone note noteGrp pb

*figures:* figure notatedMusic

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*textstructure:* back body front group

*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin

*Note* This element should not be used to represent a text which is inserted at an arbitrary point within the structure of another, for example as in an embedded or quoted narrative; the <floatingText> is provided for this purpose.

*Example*

```
<text>
  <front>
    <docTitle>
      <titlePart>Autumn Haze</titlePart>
    </docTitle>
  </front>
  <body>
    <l>Is it a dragonfly or a maple leaf</l>
    <l>That settles softly down upon the water?</l>
  </body>
</text>
```

*Example* The body of a text may be replaced by a group of nested texts, as in the following schematic:

```
<text>
  <front>
```

```
<!-- front matter for the whole group -->
  </front>
  <group>
    <text>
<!-- first text -->
    </text>
    <text>
<!-- second text -->
    </text>
  </group>
</text>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.global"
   minOccurs="0" maxOccurs="unbounded"/>
  <sequence minOccurs="0">
   <elementRef key="front"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <alternate>
   <elementRef key="body"/>
   <elementRef key="group"/>
  </alternate>
  <classRef key="model.global"
   minOccurs="0" maxOccurs="unbounded"/>
  <sequence minOccurs="0">
   <elementRef key="back"/>
   <classRef key="model.global"
    minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element text
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declaring.attributes,
   att.typed.attributes,
   att.written.attributes,
   (
      model.global*,
      ( ( front, model.global* )? ),
      ( body | group ),
      model.global*,
```

```
      ( ( back, model.global* )? )
   )
}
```

---

## <textClass>

(text classification) groups information which describes the nature or topic of a text in terms of a standard classification scheme, thesaurus, etc. [*2.4.3. The Text Classification*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable
  - *@default*

*Member of* model.profileDescPart

*Contained by*

*header:* profileDesc
*May contain*
*header:* catRef classCode keywords
*Example*

```
<taxonomy>
  <category  xml:id="acprose">
    <catDesc>Academic prose</catDesc>
  </category>
<!-- other categories here -->
</taxonomy>
<!-- ... -->
<textClass>
  <catRef  target="#acprose"/>
  <classCode  scheme="http://www.udcc.org">001.9</classCode>
  <keywords  scheme="http://authorities.loc.gov">
    <list>
      <item>End of the world</item>
      <item>History - philosophy</item>
    </list>
  </keywords>
</textClass>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:textClass"/> </sch:pattern>
*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <elementRef key="classCode"/>
  <elementRef key="catRef"/>
  <elementRef key="keywords"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element textClass
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   ( classCode | catRef | keywords )*
}
```

---

**<textDesc>** (text description) provides a description of a text in terms of its situational parameters. [*16.2.1. The Text Description*]
*Module* corpus — Language Corpora

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable
  - *@default*

*Member of* model.catDescPart model.profileDescPart

*Contained by*

*header:* catDesc profileDesc

*May contain*

*corpus:* channel constitution derivation domain factuality interaction preparedness purpose

*Example*

```
<textDesc n="Informal domestic conversation">
  <channel mode="s"/>
  <constitution type="single"/>
  <derivation type="original"/>
```

```
  <domain type="domestic"/>
  <factuality type="mixed"/>
  <interaction type="complete"
   active="plural" passive="many"/>
  <preparedness type="spontaneous"/>
  <purpose type="entertain" degree="high"/>
  <purpose type="inform" degree="medium"/>
 </textDesc>
```

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:textDesc"/> </sch:pattern>

*Content model*

```
<content>
 <sequence>
  <classRef key="model.textDescPart"
   expand="sequence"/>
  <elementRef key="purpose" minOccurs="1"
   maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element textDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   (
      channel,
      constitution,
      derivation,
      domain,
      factuality,
      interaction,
      preparedness,
      purpose+
   )
}
```

---

## <textLang>

(text language) describes the languages and writing systems identified within the bibliographic work being described, rather than its description. [*3.12.2.4. Imprint, Size of a Document, and Reprint Information 11.6.6. Languages and Writing Systems*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@mainLang` (main language) supplies a code which identifies the chief language used in the bibliographic work.
*Status* Optional
*Datatype* teidata.language

`@otherLangs` (other languages) one or more codes identifying any other languages used in the bibliographic work.
*Status* Optional
*Datatype* 0–∞ occurrences of teidata.language separated by whitespace

*Member of* model.biblPart model.msItemPart

*Contained by*

*core:* analytic bibl monogr series

*msdescription:* msContents msItem msItemStruct

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media

mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* This element should not be used to document the languages or writing systems used for the bibliographic or manuscript description itself: as for all other TEI elements, such information should be provided by means of the global *xml:lang* attribute attached to the element containing the description.

In all cases, languages should be identified by means of a standardized 'language tag' generated according to BCP 47. Additional documentation for the language may be provided by a <language> element in the TEI header.

*Example*

```
<textLang mainLang="en" otherLangs="la"> Predominantly in English with Latin
glosses</textLang>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element textLang
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
```

```
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute mainLang { text }?,
   attribute otherLangs { list { * } }?,
   macro.specialPara}
```

---

**<textNode>** indicates the presence of a text node in a content model. [*23. Documentation Elements*]

*Module* tagdocs — Documentation Elements

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Member of* model.contentPart

*Contained by*

*tagdocs:* alternate content sequence

*May contain* Empty element

*Example*

```
<content>
  <alternate  minOccurs="0"
    maxOccurs="unbounded">
    <textNode/>
    <elementRef  key="hi"/>
    <elementRef  key="seg"/>
  </alternate>
</content>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element textNode
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   empty
}
```

---

**<then>** separates the condition from the default in an <if>, or the antecedent and the consequent in a <cond> element. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*

* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Contained by* cond if
*May contain* Empty element
*Note* This element is provided primarily to enhance the human readability of the feature-system declaration.
*Example*

```
<cond>
  <fs>
    <f  name="BAR">
      <symbol  value="1"/>
    </f>
  </fs>
  <then/>
  <fs>
    <f  name="FOO">
      <binary  value="false"/>
    </f>
  </fs>
</cond>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element then
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   empty
}
```

---

## <time>

(time) contains a phrase defining a time of day in any format. [*3.6.4. Dates and Times*]
*Module* core — Elements Available in All TEI Documents

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.calendarSystem
  - *@calendar*
- att.canonical
  - *@key*
  - *@ref*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    - *@when-custom*

    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.duration
  - att.duration.iso
    * *@dur-iso*
  - att.duration.w3c
    * *@dur*
- att.editLike
  - *@evidence*
  - *@instant*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.dateLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose setting

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* recording u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q ref reg rs ruby sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat

geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
As he sat smiling, the
quarter struck — <time  when="11:45:00">the quarter to twelve</time>.
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element time
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.calendarSystem.attributes,
   att.canonical.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   att.editLike.attributes,
   att.typed.attributes,
   ( text | model.gLike | model.phrase | model.global )*
}
```

**<timeline>** (timeline) provides a set of ordered points in time which can be linked to elements of a spoken text to create a temporal alignment of that text. [*17.4.2. Placing Synchronous Events in Time*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*

`@origin` designates the origin of the timeline, i.e. the time at which it begins.
*Status* Optional
*Datatype* teidata.pointer
*Note* If this attribute is not supplied, the implication is that the time of origin is not known. If it is supplied, it must point either to one of the <when> elements in its content, or to another <timeline> element.

`@unit` specifies the unit of time corresponding to the *interval* value of the timeline or of its constituent points in time.
*Status* Optional

*Datatype* teidata.enumerated

*Suggested values include:* **d** (days)

**h** (hours)

**min** (minutes)

**s** (seconds)

**ms** (milliseconds)

`@interval` specifies a time interval either as a positive integral value or using one of a set of predefined codes.

*Status* Optional

*Datatype* teidata.interval

*Note* The value irregular indicates uncertainty about all the intervals in the timeline; the value regular indicates that all the intervals are evenly spaced, but the size of the intervals is not known; numeric values indicate evenly spaced values of the size specified. If individual points in time in the timeline are given different values for the *interval* attribute, those values locally override the value given in the timeline.

*Member of* model.global.meta

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg standOff

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme
*May contain*
*linking:* when
*Example*

```
<timeline xml:id="TL01" unit="ms">
  <when xml:id="TL-w0" absolute="11:30:00"/>
  <when xml:id="TL-w1" interval="unknown"
    since="#TL-w0"/>
  <when xml:id="TL-w2" interval="100"
    since="#TL-w1"/>
  <when xml:id="TL-w3" interval="200"
    since="#TL-w2"/>
  <when xml:id="TL-w4" interval="150"
    since="#TL-w3"/>
  <when xml:id="TL-w5" interval="250"
    since="#TL-w4"/>
  <when xml:id="TL-w6" interval="100"
    since="#TL-w5"/>
</timeline>
```

*Content model*

```
<content>
 <elementRef key="when" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element timeline
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   attribute origin { text }?,
   attribute unit { "d" | "h" | "min" | "s" | "ms" }?,
   attribute interval { text }?,
   when+
}
```

---

## <title>

(title) contains a title for any kind of work. [*3.12.2.2. Titles, Authors, and Editors 2.2.1. The Title Statement 2.2.5. The Series Statement*]

*Module* core — Elements Available in All TEI Documents
*Attributes*
- att.global
  - *@xml:id*
  - *@n*

  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    * *@ana*
  - att.global.change
    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*

  * *@datingMethod*
  – att.datable.iso
  * *@when-iso*
  * *@notBefore-iso*
  * *@notAfter-iso*
  * *@from-iso*
  * *@to-iso*
  – att.datable.w3c
  * *@when*
  * *@notBefore*
  * *@notAfter*
  * *@from*
  * *@to*
- att.typed
  – ~~type~~
  – @subtype

`@type` classifies the title according to some convenient typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Sample values include:* **main** main title
**sub** (subordinate) subtitle, title of part
**alt** (alternate) alternate title, often in another language, by which the work is also known
**short** abbreviated form of title
**desc** (descriptive) descriptive paraphrase of the work functioning as a title
*Note* This attribute is provided for convenience in analysing titles and processing them according to their type; where such specialized processing is not necessary, there is no need for such analysis, and the entire title, including subtitles and any parallel titles, may be enclosed within a single <title> element.

`@level` indicates the bibliographic level for a title, that is, whether it identifies an article, book, journal, series, or unpublished material.
*Status* Optional
*Datatype* teidata.enumerated
*Legal values are:* **a** (analytic) the title applies to an analytic item, such as an article, poem, or other work published as part of a larger item.
**m** (monographic) the title applies to a monograph such as a book or other item considered to be a distinct publication, including single volumes of multi-volume works
**j** (journal) the title applies to any serial or periodical publication such as a journal, magazine, or newspaper
**s** (series) the title applies to a series of otherwise distinct publications such as a collection
**u** (unpublished) the title applies to any unpublished material (including theses and dissertations unless published by a commercial press)
*Note* The level of a title is sometimes implied by its context: for example, a title appearing directly within an <analytic> element is *ipso facto* of level 'a', and one appearing

within a <series> element of level 's'. For this reason, the *level* attribute is not required in contexts where its value can be unambiguously inferred. Where it is supplied in such contexts, its value should not contradict the value implied by its parent element.

*Member of* model.emphLike model.msQuoteLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine analytic author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned monogr name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote seriesStmt sponsor tagUsage titleStmt

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem msItemStruct musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The attributes *key* and *ref*, inherited from the class `att.canonical` may be used to indicate the canonical form for the title; the former, by supplying (for example) the identifier of a record in some external library system; the latter by pointing to an XML element somewhere containing the canonical form of the title.

*Example*

```
<title>Information Technology and the Research Process: Proceedings of
a conference held at Cranfield Institute of Technology, UK,
18–21 July 1989</title>
```

*Example*

```
<title>Hardy's Tess of the D'Urbervilles: a machine readable
edition</title>
```

*Example*

```
<title  type="full">
  <title  type="main">Synthèse</title>
  <title  type="sub">an international journal for
     epistemology, methodology and history of
     science</title>
</title>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element title
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
```

```
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   attribute level { "a" | "m" | "j" | "s" | "u" }?,
   macro.paraContent}
```

---

## <titlePage>

(title page) contains the title page of a text, appearing within the front or back matter. [*4.6. Title Pages*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

* *@cert*
* *@resp*

– att.global.source
* *@source*

• att.typed
– ~~type~~
– @subtype

`@type` classifies the title page according to any convenient typology.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Note* This attribute allows the same element to be used for volume title pages, series title pages, etc., as well as for the 'main' title page of a work.

*Member of* model.frontPart
*Contained by*
*msdescription:* msContents
*textcrit:* lem rdg
*textstructure:* back front
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* binaryObject cb ellipsis gap gb graphic index lb milestone note noteGrp pb
*figures:* figure notatedMusic
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline
*spoken:* incident kinesic pause shift vocal writing
*textcrit:* app witDetail
*textstructure:* argument byline docAuthor docDate docEdition docImprint docTitle epigraph imprimatur titlePart
*transcr:* addSpan damageSpan delSpan fw listTranspose metamark space substJoin
*Example*

```
<titlePage>
  <docTitle>
    <titlePart  type="main">THOMAS OF Reading.</titlePart>
    <titlePart  type="alt">OR, The sixe worthy yeomen of the West.</titlePart>
  </docTitle>
  <docEdition>Now the fourth time corrected and enlarged</docEdition>
  <byline>By T.D.</byline>
  <figure>
    <head>TP</head>
    <p>Thou shalt labor till thou returne to duste</p>
    <figDesc>Printers Ornament used by TP</figDesc>
  </figure>
  <docImprint>Printed at <name  type="place">London</name> for
<name>T.P.</name>
    <date>1612.</date>
  </docImprint>
</titlePage>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.global"
   minOccurs="0" maxOccurs="unbounded"/>
  <classRef key="model.titlepagePart"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <classRef key="model.titlepagePart"/>
   <classRef key="model.global"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element titlePage
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   (
      model.global*,
      model.titlepagePart,
      ( model.titlepagePart | model.global )*
   )
}
```

---

## <titlePart>

<titlePart> (title part) contains a subsection or division of the title of a work, as indicated on a title page. [*4.6. Title Pages*]

*Module* textstructure — Default Text Structure

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs

    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.typed
  - ~~type~~
  - @subtype

`@type` (type) specifies the role of this subdivision of the title.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **main** (main) main title of the work[Default]
**sub** (subordinate) subtitle of the work
**alt** (alternate) alternative title of the work
**short** (short) abbreviated form of title
**desc** (descriptive) descriptive paraphrase of the work

*Member of* model.pLike.front model.titlepagePart
*Contained by*
*msdescription:* msItem
*textcrit:* lem rdg
*textstructure:* back docTitle front titlePage
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<docTitle>
  <titlePart  type="main">THE FORTUNES
     AND MISFORTUNES Of the FAMOUS
     Moll Flanders, &c.
  </titlePart>
  <titlePart  type="desc">Who was BORN in NEWGATE,
     And during a Life of continu'd Variety for
     Threescore Years, besides her Childhood, was
     Twelve Year a <hi>Whore</hi>, five times a <hi>Wife</hi> (wherof
     once to her own Brother) Twelve Year a <hi>Thief,</hi>
     Eight Year a Transported <hi>Felon</hi> in <hi>Virginia</hi>,
     at last grew <hi>Rich</hi>, liv'd <hi>Honest</hi>, and died a
  <hi>Penitent</hi>.</titlePart>
</docTitle>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element titlePart
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
att.global.source.attributes,
att.typed.attribute.subtype,
attribute type { "main" | "sub" | "alt" | "short" | "desc" }?,
macro.paraContent}
```

---

**<titleStmt>** (title statement) groups information about the title of a work and those responsible for its content.
[*2.2.1. The Title Statement 2.2. The File Description*]

*Module* header — The TEI Header

*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  - * *@ana*
- – att.global.change
  - * *@change*
- – att.global.facs
  - * *@facs*
- – att.global.linking
  - * *@corresp*
  - * *@synch*
  - * *@sameAs*
  - * *@copyOf*
  - * *@next*
  - * *@prev*
  - * *@exclude*
  - * *@select*
- – att.global.rendition
  - * *@rend*
  - * *@style*
  - * *@rendition*
- – att.global.responsibility
  - * *@cert*
  - * *@resp*
- – att.global.source
  - * *@source*

*Contained by:* header: biblFull fileDesc

*May contain*

*core:* author editor meeting respStmt title

*header:* funder principal sponsor
*Example*

```
<titleStmt>
  <title>Capgrave's Life of St. John Norbert: a machine-readable
transcription</title>
  <respStmt>
    <resp>compiled by</resp>
    <name>P.J. Lucas</name>
  </respStmt>
</titleStmt>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="title" minOccurs="1"
   maxOccurs="unbounded"/>
  <classRef key="model.respLike"
   minOccurs="0" maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element titleStmt
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( title+, model.respLike* )
}
```

---

**<tns>** (tense) indicates the grammatical tense associated with a given inflected form in a dictionary. [*10.3.1. Information on Written and Spoken Forms*]

*Module* dictionaries — Dictionaries
*Attributes* • att.global
- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  * *@ana*
- att.global.change

    * *@change*
  - att.global.facs
    * *@facs*
  - att.global.linking
    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.lexicographic.normalized
    * *@norm*
    * *@orig*

*Member of* model.morphLike
*Contained by*
*dictionaries:* dictScrap entryFree form gramGrp
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph

expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* This element is synonymous with <gram type="tense">.

*Example* Taken from *Wörterbuch der Deutschen Sprache. Veranstaltet und herausgegeben von Joachim Heinrich Campe. Vierter Theil. S - bis - T. (Braunschweig 1810. In der Schulbuchhandlung)*:

> Treffen, v. unregelm. … du triffst, …

```
<entry>
  <form type="inflected">
    <gramGrp>
      <per value="2"/>
      <number value="singular"/>
      <tns value="present"/>
      <mood value="indicative"/>
    </gramGrp>
    <form type="personalpronoun">
      <orth>du</orth>
    </form>
    <form type="headword">
      <orth>
        <oRef>triffst</oRef>
      </orth>
    </form>
  </form>
</entry>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element tns
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   macro.paraContent}
```

---

**<trailer>** contains a closing title or footer appearing at the end of a division of a text. [*4.2.4. Content of Textual Divisions 4.2. Elements Common to All Divisions*]

*Module* textstructure — Default Text Structure

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.placement
  - *@place*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*

*Member of* model.divBottomPart

*Contained by*

*cmc:* post

*core:* lg list

*drama:* castGroup epilogue performance prologue

*figures:* figure table

*textstructure:* back body div div1 div2 div3 div4 div5 div6 div7 front group postscript

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<trailer>Explicit pars tertia</trailer>
```

*Example*

```
<trailer>
  <l>In stead of FINIS this advice <hi>I</hi> send,</l>
  <l>Let Rogues and Thieves beware of <lb/>
    <hi>Hamans</hi> END.</l>
</trailer>
```

From EEBO A87070

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <elementRef key="lg"/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.lLike"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element trailer
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.placement.attributes,
   att.typed.attributes,
   att.written.attributes,
   (
      text
```

```
    | lg    | model.gLike    | model.phrase    | model.inter    | model.lLike    | model.g
}
```

---

**<trait>** (trait) contains a description of some status or quality attributed to a person, place, or organization typically, but not necessarily, independent of the volition or action of the holder and usually not at some specific time or for a specific date range. [*14.3.1. Basic Principles 14.3.2.1. Personal Characteristics*]

*Module* namesdates — Names, Dates, People, and Places

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*
- att.editLike
  - *@evidence*
  - *@instant*
- att.naming
  - *@role*
  - *@nymRef*

  - att.canonical
    * *@key*

    * *@ref*
  - att.typed
    - *@type*
    - *@subtype*

*Member of* model.persStateLike model.placeStateLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode correspAction creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset org orgName persName persPronouns person personGrp persona place placeName region residence roleName settlement sex socecStatus surname trait

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*certainty:* precision

*core:* bibl biblStruct desc head label listBibl note noteGrp p

*header:* biblFull

*linking:* ab

*msdescription:* msDesc

*namesdates:* trait

*Note* Where there is confusion between <trait> and <state> the more general purpose element <state> should be used even for unchanging characteristics. If you wish to distinguish between characteristics that are generally perceived to be time-bound states and those assumed to be fixed traits, then <trait> is available for the more static of these. The <state> element encodes characteristics which are sometimes assumed to change, often at specific times or over a date range, whereas the <trait> elements are used to record

characteristics, such as eye-colour, which are less subject to change. Traits are typically, but not necessarily, independent of the volition or action of the holder.

*Example*

```
<trait  type="physical">
  <label>Eye colour</label>
  <desc>Blue</desc>
</trait>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="precision" minOccurs="0"
   maxOccurs="unbounded"/>
  <alternate>
   <elementRef key="trait" minOccurs="1"
    maxOccurs="unbounded"/>
   <sequence>
    <classRef key="model.headLike"
     minOccurs="0" maxOccurs="unbounded"/>
    <classRef key="model.pLike"
     minOccurs="1" maxOccurs="unbounded"/>
    <alternate minOccurs="0"
     maxOccurs="unbounded">
     <classRef key="model.noteLike"/>
     <classRef key="model.biblLike"/>
    </alternate>
   </sequence>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.labelLike"/>
    <classRef key="model.noteLike"/>
    <classRef key="model.biblLike"/>
   </alternate>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element trait
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
```

```
   att.editLike.attributes,
   att.naming.attributes,
   att.canonical.attributes,
   att.typed.attributes,
   (
      precision*,
      (
         trait+
       | (
            model.headLike*,
            model.pLike+,
            ( model.noteLike | model.biblLike )*
         )
       | ( model.labelLike | model.noteLike | model.biblLike )*
      )
   )
}
```

---

## <transcriptionDesc>

(transcription description) describes the set of transcription conventions used, particularly for spoken material. [*8.2. Documenting the Source of Transcribed Speech*]

*Module* spoken — Transcriptions of Speech

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*

* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@ident` supplies an identifier for the encoding convention, independent of any version number.
*Status* Required
*Datatype* teidata.name

`@version` supplies a version number for the encoding conventions used, if any.
*Status* Optional
*Datatype* teidata.versionNumber

*Member of* model.encodingDescPart
*Contained by*
*header:* encodingDesc
*May contain*
*core:* desc label p ptr ref
*linking:* ab
*tagdocs:* listRef
*Example*

```
<transcriptionDesc ident="HIAT"
  version="2004"/>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <classRef key="model.labelLike"/>
  <classRef key="model.ptrLike"/>
  <classRef key="model.pLike"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element transcriptionDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute ident { text },
   attribute version { text }?,
   ( model.labelLike | model.ptrLike | model.pLike )*
}
```

**<transpose>** describes a single textual transposition as an ordered list of at least two pointers specifying the order in which the elements indicated should be re-combined. [*12.3.4.5. Transpositions*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

*Contained by* transcr: listTranspose

*May contain*

*core:* ptr

*Note* Transposition is usually indicated in a document by a metamark such as a wavy line or numbering.

The order in which <ptr> elements appear within a <transpose> element should correspond with the desired order, as indicated by the metamark.

*Example*

```
<transpose>
  <ptr  target="#ib02"/>
  <ptr  target="#ib01"/>
</transpose>
```

The transposition recorded here indicates that the content of the element with identifier `ib02` should appear before the content of the element with identifier `ib01`.

*Content model*

```
<content>
 <elementRef key="ptr" minOccurs="2"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element transpose
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( ptr, ptr, ptr* )
}
```

---

**<tree>** (tree) encodes a tree, which is made up of a root, internal nodes, leaves, and arcs from root to leaves. [*20.2. Trees*]

*Module* nets — Graphs, Networks, and Trees

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*

  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- att.global.responsibility
  * *@cert*
  * *@resp*
- att.global.source
  * *@source*

`@arity` gives the maximum number of children of the root and internal nodes of the tree.
*Status* Optional
*Datatype* teidata.count

`@ord` (ordered) indicates whether or not the tree is ordered, or if it is partially ordered.
*Status* Required
*Datatype* teidata.enumerated
*Legal values are:* **true** indicates that all of the branching nodes of the tree are ordered.
**partial** indicates that some of the branching nodes of the tree are ordered and some are unordered.
**false** indicates that all of the branching nodes of the tree are unordered.

`@order` gives the order of the tree, i.e., the number of its nodes.
*Status* Optional
*Datatype* teidata.count
*Note* The size of a tree is always one less than its order, hence there is no need for both a *size* and *order* attribute.

*Member of* model.divPart model.standOffPart
*Contained by*
*core:* item note q quote said stage textLang
*drama:* castList epilogue performance prologue set view
*figures:* cell figure
*header:* change handNote licence scriptNote
*linking:* standOff
*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote
*namesdates:* occupation
*nets:* forest
*tagdocs:* specGrp
*textcrit:* lem rdg
*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript
*transcr:* metamark

*May contain*

*core:* label

*nets:* iNode leaf root

*Note* A root, and zero or more internal nodes and leaves, but if there is an internal node, there must also be at least one leaf.

*Example*

```
<tree n="ex2" arity="2" ord="partial"
  order="13">
  <root xml:id="G-div1"
    children="#G-plu1 #G-exp1" ord="true">
    <label>/</label>
  </root>
  <iNode xml:id="G-plu1"
    children="#G-exp2 #G-exp3" parent="#G-div1" ord="false">
    <label>+</label>
  </iNode>
  <iNode xml:id="G-exp1"
    children="#G-plu2 #G-num2.3" parent="#G-div1" ord="true">
    <label>**</label>
  </iNode>
  <iNode xml:id="G-exp2"
    children="#G-vara1 #G-num2.1" parent="#G-plu1" ord="true">
    <label>**</label>
  </iNode>
  <iNode xml:id="G-exp3"
    children="#G-varb1 #G-num2.2" parent="#G-plu1" ord="true">
    <label>**</label>
  </iNode>
  <iNode xml:id="G-plu2"
    children="#G-vara2 #G-varb2" parent="#G-exp1" ord="false">
    <label>+</label>
  </iNode>
  <leaf xml:id="G-vara1" parent="#G-exp2">
    <label>a</label>
  </leaf>
  <leaf xml:id="G-num2.1" parent="#G-exp2">
    <label>2</label>
  </leaf>
  <leaf xml:id="G-varb1" parent="#G-exp3">
    <label>b</label>
  </leaf>
  <leaf xml:id="G-num2.2" parent="#G-exp3">
    <label>2</label>
  </leaf>
  <leaf xml:id="G-vara2" parent="#G-plu2">
    <label>a</label>
  </leaf>
  <leaf xml:id="G-varb2" parent="#G-plu2">
    <label>b</label>
  </leaf>
  <leaf xml:id="G-num2.3" parent="#G-exp1">
    <label>2</label>
  </leaf>
</tree>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="label" minOccurs="0"/>
  <sequence>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="leaf"/>
    <elementRef key="iNode"/>
   </alternate>
   <elementRef key="root"/>
   <alternate minOccurs="0"
    maxOccurs="unbounded">
    <elementRef key="leaf"/>
    <elementRef key="iNode"/>
   </alternate>
  </sequence>
 </sequence>
</content>
```

*Schema Declaration*

```
element tree
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute arity { text }?,
   attribute ord { "true" | "partial" | "false" },
   attribute order { text }?,
   ( label?, ( ( leaf | iNode )*, root, ( leaf | iNode )* ) )
}
```

---

**<triangle>** (underspecified embedding tree, so called because of its characteristic shape when drawn) provides for an underspecified <eTree>, that is, an <eTree> with information left out. [*20.3. Another Tree Notation*]

*Module* nets — Graphs, Networks, and Trees

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change

* *@change*

– att.global.facs
* *@facs*

– att.global.linking
* *@corresp*
* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@value` supplies a value for the triangle, in the form of the identifier of a feature structure or other analytic element.
*Status* Optional
*Datatype* teidata.pointer

*Contained by* *nets:* eTree forest triangle
*May contain*
*core:* label
*nets:* eLeaf eTree triangle
*Note* An optional label followed by zero or more embedding trees, triangles, or embedding leafs.
*Example*

```
<triangle>
  <label>NP</label>
  <eLeaf>
    <label>the periscope</label>
  </eLeaf>
</triangle>
```

*Content model*

```
<content>
 <sequence>
  <elementRef key="label" minOccurs="0"/>
  <alternate minOccurs="0"
   maxOccurs="unbounded">
   <elementRef key="eTree"/>
```

```
   <elementRef key="triangle"/>
   <elementRef key="eLeaf"/>
  </alternate>
 </sequence>
</content>
```

*Schema Declaration*

```
element triangle
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute value { text }?,
   ( label?, ( eTree | triangle | eLeaf )* )
}
```

---

## <typeDesc>

<typeDesc> (typeface description) contains a description of the typefaces or other aspects of the printing of an incunable or other printed source. [*11.7.2.1. Writing*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*

* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Member of* model.physDescPart
*Contained by*
*msdescription:* physDesc
*May contain*
*core:* p
*linking:* ab
*msdescription:* summary typeNote
*Example*

```
<typeDesc>
  <p>Uses an unidentified black letter font, probably from the
     15th century</p>
</typeDesc>
```

*Example*

```
<typeDesc>
  <summary>Contains a mixture of blackletter and Roman (antiqua)
typefaces</summary>
  <typeNote  xml:id="Frak1">Blackletter face, showing
     similarities to those produced in Wuerzburg after 1470.</typeNote>
  <typeNote  xml:id="Rom1">Roman face of Venetian origins.</typeNote>
</typeDesc>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.pLike" minOccurs="1"
   maxOccurs="unbounded"/>
  <sequence>
   <elementRef key="summary" minOccurs="0"/>
   <elementRef key="typeNote" minOccurs="1"
    maxOccurs="unbounded"/>
  </sequence>
 </alternate>
</content>
```

*Schema Declaration*

```
element typeDesc
{
```

```
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.pLike+ | ( summary?, typeNote+ ) )
}
```

---

## <typeNote>

(typographic note) describes a particular font or other significant typographic feature distinguished within the description of a printed resource. [*11.7.2. Writing, Decoration, and Other Notations*]

*Module* msdescription — Manuscript Description

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

    * *@source*

  - att.handFeatures
    – *@scribe*
    – *@scribeRef*
    – *@script*
    – *@scriptRef*
    – *@medium*

    – att.scope
    * *@scope*

*Contained by* *msdescription:* typeDesc
*May contain*
*analysis:* c cl interp interpGrp m pc phr s span spanGrp w
*certainty:* certainty precision respons
*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit
*dictionaries:* lang oRef pRef
*drama:* camera caption castList move sound spGrp tech view
*figures:* figure formula notatedMusic table
*gaiji:* g
*header:* biblFull idno
*iso-fs:* fLib fs fvLib
*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline
*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width
*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait
*nets:* eTree forest graph listForest tree
*spoken:* annotationBlock incident kinesic pause shift u vocal writing
*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val
*textcrit:* app listApp listWit witDetail
*textstructure:* floatingText
*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo
*verse:* caesura rhyme
character data
*Example*

```
<typeNote  scope="sole"> Printed in an Antiqua typeface showing strong
Italianate influence.
</typeNote>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element typeNote
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.handFeatures.attributes,
   att.scope.attributes,
   macro.specialPara}
```

---

**<u>** (utterance) contains a stretch of speech usually preceded and followed by silence or by a change of speaker. [*8.3.1. Utterances*]

*Module* spoken — Transcriptions of Speech

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*
- att.ascribed.directed
  - *@toWhom*
- att.ascribed
  - *@who*
- att.declaring
  - *@decls*
- att.notated
  - *@notation*
- att.timed
  - *@start*
  - *@end*
- att.duration
  - att.duration.iso
    - *@dur-iso*
  - att.duration.w3c
    - *@dur*

`@trans` (transition) indicates the nature of the transition between this utterance and the previous one.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **smooth** this utterance begins without unusual pause or rapidity.[Default]

**latching** this utterance begins with a markedly shorter pause than normal.

**overlap** this utterance begins before the previous one has finished.

**pause** this utterance begins after a noticeable pause.

*Member of* model.divPart.spoken model.standOffPart

*Contained by*

*core:* item note q quote said stage textLang

*drama:* castList epilogue performance prologue set view

*figures:* cell figure

*header:* change handNote licence scriptNote

*linking:* standOff

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* annotationBlock

*tagdocs:* specGrp

*textcrit:* lem rdg

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 epigraph postscript

*transcr:* metamark

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* Prose and a mixture of speech elements

Although individual transcriptions may consistently use <u> elements for turns or other units, and although in most cases a <u> will be delimited by pause or change of speaker, <u> is not required to represent a turn or any communicative event, nor to be bounded by pauses or change of speaker. At a minimum, a <u> is some phonetic production by a given speaker.

*Example*

```
<u  who="#spkr1">if did you set</u>
<u  trans="latching"  who="#spkr2">well Joe and I set it between us</u>
<list  type="speakers">
  <item  xml:id="spkr1"/>
  <item  xml:id="spkr2"/>
</list>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element u
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.directed.attributes,
   att.ascribed.attributes,
   att.declaring.attributes,
   att.notated.attributes,
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   attribute trans { "smooth" | "latching" | "overlap" | "pause" }?,
   macro.paraContent}
```

---

**<unclear>** (unclear) contains a word, phrase, or passage which cannot be transcribed with certainty because it is illegible or inaudible in the source. [*12.3.3.1. Damage, Illegibility, and Supplied Text 3.5.3. Additions, Deletions, and Omissions*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*

  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
  - att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
  - att.global.responsibility
  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.cmc
  - *@generatedBy*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
  * *@atLeast*
  * *@atMost*
  * *@min*
  * *@max*
  * *@confidence*
- att.editLike
  - *@evidence*
  - *@instant*

`@reason` indicates why the material is hard to transcribe.

*Status* Optional

*Datatype* 1–∞ occurrences of teidata.enumerated separated by whitespace

*Suggested values include:* **illegible** (illegible)

**inaudible** (inaudible)

**faded** (faded)

**background_noise** (background noise)

**eccentric_ductus** (eccentric ductus) indicates illegibility due to an unusual, awkward, or incompetent execution of a glyph or glyphs

```
<div>
  <head>Rx</head>
  <p>500 mg <unclear  reason="illegible">placebo</unclear>
```

```
    </p>
  </div>
```

*Note* One or more words may be used to describe the reason; usually each word will refer to a single cause.

`@agent` Where the difficulty in transcription arises from damage, categorizes the cause of the damage, if it can be identified.

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **rubbing** damage results from rubbing of the leaf edges

**mildew** damage results from mildew on the leaf surface

**smoke** damage results from smoke

*Member of* model.choicePart model.linePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope choice citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The same element is used for all cases of uncertainty in the transcription of element content, whether for written or spoken material. For other aspects of certainty, uncertainty, and reliability of tagging and transcription, see chapter *22. Certainty, Precision, and Responsibility*.

The <damage>, <gap>, <del>, <unclear> and <supplied> elements may be closely allied in use. See section *12.3.3.2. Use of the gap, del, damage, unclear, and supplied Elements in Combination* for discussion of which element is appropriate for which circumstance.

The *hand* attribute points to a definition of the hand concerned, as further discussed in section *12.3.2.1. Document Hands*.

*Example*

```
<u> ...and then <unclear  reason="background-noise">Nathalie</unclear> said ...
</u>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element unclear
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
```

```
   att.editLike.attributes,
   attribute reason
   {
      list
      {
         (
            "illegible"
          | "inaudible"
          | "faded"
          | "background_noise"
          | "eccentric_ductus"
         )+
      }
   }?,
   attribute agent { text }?,
   macro.paraContent}
```

---

**<undo>** indicates one or more marked-up interventions in a document which have subsequently been marked for cancellation. [*12.3.4.4. Confirmation, Cancellation, and Reinstatement of Modifications*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*

  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*
  - att.ranging
    - *@atLeast*
    - *@atMost*
    - *@min*
    - *@max*
    - *@confidence*
- att.spanning
  - *@spanTo*
- att.transcriptional
  - *@status*
  - *@cause*
  - *@seq*
  - att.editLike
    - *@evidence*
    - *@instant*
  - att.placement
    - *@place*
  - att.written
    - *@hand*

`@target` points to one or more elements representing the interventions which are to be reverted or undone.
*Status* Optional
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Member of* model.linePart model.pPart.transcriptional

*Contained by*

*analysis:* cl pc phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label lg measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* am damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain* Empty element

*Example*

```
<line>This is <del change="#s2" rend="overstrike">
    <seg xml:id="undo-a">just some</seg>
     sample <seg xml:id="undo-b">text</seg>,
     we need</del>
  <add change="#s2">not</add>
a real example.</line>
<undo target="#undo-a #undo-b"
  rend="dotted" change="#s3"/>
```

This encoding represents the following sequence of events:

- ”This is just some sample text, we need a real example” is written
- At stage s2, ”just some sample text, we need” is deleted by overstriking, and ”not” is added
- At stage s3, parts of the deletion are cancelled by underdotting, thus reinstating the words ”just some” and ”text”.

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element undo
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
```

```
   att.spanning.attributes,
   att.transcriptional.attributes,
   att.editLike.attributes,
   att.placement.attributes,
   att.written.attributes,
   attribute target { list { + } }?,
   empty
}
```

---

**<unicodeProp>** (unicode property) provides a Unicode property for a character (or glyph). [*5.2.1. Character Properties*]

*Module* gaiji — Characters, Glyphs, and Writing Modes

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.gaijiProp

- ~~name~~
- ~~value~~
- @version
- @scheme

- att.datable
  - @period
- att.datable.custom
  - @when-custom
  - @notBefore-custom
  - @notAfter-custom
  - @from-custom
  - @to-custom
  - @datingPoint
  - @datingMethod
- att.datable.iso
  - @when-iso
  - @notBefore-iso
  - @notAfter-iso
  - @from-iso
  - @to-iso
- att.datable.w3c
  - @when
  - @notBefore
  - @notAfter
  - @from
  - @to

`@name` specifies the normalized name of a Unicode property.

*Status* Required

*Datatype* teidata.xmlName

*Legal values are:* **Age**
**AHex**
**Alpha**
**Alphabetic**
**ASCII_Hex_Digit**
**bc**
**Bidi_C**
**Bidi_Class**
**Bidi_Control**
**Bidi_M**
**Bidi_Mirrored**
**Bidi_Mirroring_Glyph**
**Bidi_Paired_Bracket**
**Bidi_Paired_Bracket_Type**
**blk**

**Block**
**bmg**
**bpb**
**bpt**
**Canonical_Combining_Class**
**Case_Folding**
**Case_Ignorable**
**Cased**
**ccc**
**CE**
**cf**
**Changes_When_Casefolded**
**Changes_When_Casemapped**
**Changes_When_Lowercased**
**Changes_When_NFKC_Casefolded**
**Changes_When_Titlecased**
**Changes_When_Uppercased**
**CI**
**Comp_Ex**
**Composition_Exclusion**
**CWCF**
**CWCM**
**CWKCF**
**CWL**
**CWT**
**CWU**
**Dash**
**Decomposition_Mapping**
**Decomposition_Type**
**Default_Ignorable_Code_Point**
**Dep**
**Deprecated**
**DI**
**Dia**
**Diacritic**
**dm**
**dt**
**ea**
**East_Asian_Width**
**EqUIdeo**
**Equivalent_Unified_Ideograph**
**Expands_On_NFC**
**Expands_On_NFD**
**Expands_On_NFKC**
**Expands_On_NFKD**
**Ext**

**Extender**
**FC_NFKC**
**FC_NFKC_Closure**
**Full_Composition_Exclusion**
**gc**
**GCB**
**General_Category**
**Gr_Base**
**Gr_Ext**
**Gr_Link**
**Grapheme_Base**
**Grapheme_Cluster_Break**
**Grapheme_Extend**
**Grapheme_Link**
**Hangul_Syllable_Type**
**Hex**
**Hex_Digit**
**hst**
**Hyphen**
**ID_Continue**
**ID_Start**
**IDC**
**Ideo**
**Ideographic**
**IDS**
**IDS_Binary_Operator**
**IDS_Trinary_Operator**
**IDSB**
**IDST**
**Indic_Positional_Category**
**Indic_Syllabic_Category**
**InPC**
**InSC**
**isc**
**ISO_Comment**
**Jamo_Short_Name**
**jg**
**Join_C**
**Join_Control**
**Joining_Group**
**Joining_Type**
**JSN**
**jt**
**kAccountingNumeric**
**kCompatibilityVariant**
**kIICore**

kIRG_GSource
kIRG_HSource
kIRG_JSource
kIRG_KPSource
kIRG_KSource
kIRG_MSource
kIRG_TSource
kIRG_USource
kIRG_VSource
kOtherNumeric
kPrimaryNumeric
kRSUnicode
lb
lc
Line_Break
LOE
Logical_Order_Exception
Lower
Lowercase
Lowercase_Mapping
Math
na
na1
Name
Name_Alias
NChar
NFC_QC
NFC_Quick_Check
NFD_QC
NFD_Quick_Check
NFKC_Casefold
NFKC_CF
NFKC_QC
NFKC_Quick_Check
NFKD_QC
NFKD_Quick_Check
Noncharacter_Code_Point
nt
Numeric_Type
Numeric_Value
nv
OAlpha
ODI
OGr_Ext
OIDC
OIDS

**OLower**
**OMath**
**Other_Alphabetic**
**Other_Default_Ignorable_Code_Point**
**Other_Grapheme_Extend**
**Other_ID_Continue**
**Other_ID_Start**
**Other_Lowercase**
**Other_Math**
**Other_Uppercase**
**OUpper**
**Pat_Syn**
**Pat_WS**
**Pattern_Syntax**
**Pattern_White_Space**
**PCM**
**Prepended_Concatenation_Mark**
**QMark**
**Quotation_Mark**
**Radical**
**Regional_Indicator**
**RI**
**SB**
**sc**
**scf**
**Script**
**Script_Extensions**
**scx**
**SD**
**Sentence_Break**
**Sentence_Terminal**
**Simple_Case_Folding**
**Simple_Lowercase_Mapping**
**Simple_Titlecase_Mapping**
**Simple_Uppercase_Mapping**
**slc**
**Soft_Dotted**
**stc**
**STerm**
**suc**
**tc**
**Term**
**Terminal_Punctuation**
**Titlecase_Mapping**
**uc**
**UIdeo**

**Unicode_1_Name**
**Unified_Ideograph**
**Upper**
**Uppercase**
**Uppercase_Mapping**
**Variation_Selector**
**Vertical_Orientation**
**vo**
**VS**
**WB**
**White_Space**
**Word_Break**
**WSpace**
**XID_Continue**
**XID_Start**
**XIDC**
**XIDS**
**XO_NFC**
**XO_NFD**
**XO_NFKC**
**XO_NFKD**

`@value` specifies the value of a named Unicode property.
*Status* Required
*Datatype* teidata.text

*Containing div* char glyph
*May contain* Empty element
*Note* A definitive list of current Unicode property names is provided in The Unicode Standard.
*Example*

```
<char xml:id="U4EBA_circled">
  <unicodeProp name="Decomposition_Mapping"
    value="circle" version="12.1"/>
  <localProp name="Name"
    value="CIRCLED IDEOGRAPH 4EBA"/>
  <localProp name="daikanwa" value="36"/>
  <mapping type="standard">人</mapping>
</char>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element unicodeProp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   att.gaijiProp.attribute.version,
   att.gaijiProp.attribute.scheme,
   att.datable.attribute.period,
   att.datable.custom.attribute.when-custom,
   att.datable.custom.attribute.notBefore-custom,
   att.datable.custom.attribute.notAfter-custom,
   att.datable.custom.attribute.from-custom,
   att.datable.custom.attribute.to-custom,
   att.datable.custom.attribute.datingPoint,
   att.datable.custom.attribute.datingMethod,
   att.datable.iso.attribute.when-iso,
   att.datable.iso.attribute.notBefore-iso,
   att.datable.iso.attribute.notAfter-iso,
   att.datable.iso.attribute.from-iso,
   att.datable.iso.attribute.to-iso,
   att.datable.w3c.attribute.when,
   att.datable.w3c.attribute.notBefore,
   att.datable.w3c.attribute.notAfter,
   att.datable.w3c.attribute.from,
   att.datable.w3c.attribute.to,
   attribute name
   {
      "Age"
    | "AHex"
    | "Alpha"
    | "Alphabetic"
    | "ASCII_Hex_Digit"
    | "bc"
    | "Bidi_C"
    | "Bidi_Class"
    | "Bidi_Control"
    | "Bidi_M"
    | "Bidi_Mirrored"
    | "Bidi_Mirroring_Glyph"
    | "Bidi_Paired_Bracket"
    | "Bidi_Paired_Bracket_Type"
    | "blk"
    | "Block"
    | "bmg"
    | "bpb"
    | "bpt"
    | "Canonical_Combining_Class"
    | "Case_Folding"
    | "Case_Ignorable"
    | "Cased"
    | "ccc"
    | "CE"
    | "cf"
    | "Changes_When_Casefolded"
    | "Changes_When_Casemapped"
    | "Changes_When_Lowercased"
    | "Changes_When_NFKC_Casefolded"
    | "Changes_When_Titlecased"
    | "Changes_When_Uppercased"
    | "CI"
    | "Comp_Ex"
    | "Composition_Exclusion"
    | "CWCF"
```

```
    | "CWCM"
    | "CWKCF"
    | "CWL"
    | "CWT"
    | "CWU"
    | "Dash"
    | "Decomposition_Mapping"
    | "Decomposition_Type"
    | "Default_Ignorable_Code_Point"
    | "Dep"
    | "Deprecated"
    | "DI"
    | "Dia"
    | "Diacritic"
    | "dm"
    | "dt"
    | "ea"
    | "East_Asian_Width"
    | "EqUIdeo"
    | "Equivalent_Unified_Ideograph"
    | "Expands_On_NFC"
    | "Expands_On_NFD"
    | "Expands_On_NFKC"
    | "Expands_On_NFKD"
    | "Ext"
    | "Extender"
    | "FC_NFKC"
    | "FC_NFKC_Closure"
    | "Full_Composition_Exclusion"
    | "gc"
    | "GCB"
    | "General_Category"
    | "Gr_Base"
    | "Gr_Ext"
    | "Gr_Link"
    | "Grapheme_Base"
    | "Grapheme_Cluster_Break"
    | "Grapheme_Extend"
    | "Grapheme_Link"
    | "Hangul_Syllable_Type"
    | "Hex"
    | "Hex_Digit"
    | "hst"
    | "Hyphen"
    | "ID_Continue"
    | "ID_Start"
    | "IDC"
    | "Ideo"
    | "Ideographic"
    | "IDS"
    | "IDS_Binary_Operator"
    | "IDS_Trinary_Operator"
    | "IDSB"
    | "IDST"
    | "Indic_Positional_Category"
    | "Indic_Syllabic_Category"
    | "InPC"
    | "InSC"
    | "isc"
```

```
    | "ISO_Comment"
    | "Jamo_Short_Name"
    | "jg"
    | "Join_C"
    | "Join_Control"
    | "Joining_Group"
    | "Joining_Type"
    | "JSN"
    | "jt"
    | "kAccountingNumeric"
    | "kCompatibilityVariant"
    | "kIICore"
    | "kIRG_GSource"
    | "kIRG_HSource"
    | "kIRG_JSource"
    | "kIRG_KPSource"
    | "kIRG_KSource"
    | "kIRG_MSource"
    | "kIRG_TSource"
    | "kIRG_USource"
    | "kIRG_VSource"
    | "kOtherNumeric"
    | "kPrimaryNumeric"
    | "kRSUnicode"
    | "lb"
    | "lc"
    | "Line_Break"
    | "LOE"
    | "Logical_Order_Exception"
    | "Lower"
    | "Lowercase"
    | "Lowercase_Mapping"
    | "Math"
    | "na"
    | "na1"
    | "Name"
    | "Name_Alias"
    | "NChar"
    | "NFC_QC"
    | "NFC_Quick_Check"
    | "NFD_QC"
    | "NFD_Quick_Check"
    | "NFKC_Casefold"
    | "NFKC_CF"
    | "NFKC_QC"
    | "NFKC_Quick_Check"
    | "NFKD_QC"
    | "NFKD_Quick_Check"
    | "Noncharacter_Code_Point"
    | "nt"
    | "Numeric_Type"
    | "Numeric_Value"
    | "nv"
    | "OAlpha"
    | "ODI"
    | "OGr_Ext"
    | "OIDC"
    | "OIDS"
    | "OLower"
```

```
    | "OMath"
    | "Other_Alphabetic"
    | "Other_Default_Ignorable_Code_Point"
    | "Other_Grapheme_Extend"
    | "Other_ID_Continue"
    | "Other_ID_Start"
    | "Other_Lowercase"
    | "Other_Math"
    | "Other_Uppercase"
    | "OUpper"
    | "Pat_Syn"
    | "Pat_WS"
    | "Pattern_Syntax"
    | "Pattern_White_Space"
    | "PCM"
    | "Prepended_Concatenation_Mark"
    | "QMark"
    | "Quotation_Mark"
    | "Radical"
    | "Regional_Indicator"
    | "RI"
    | "SB"
    | "sc"
    | "scf"
    | "Script"
    | "Script_Extensions"
    | "scx"
    | "SD"
    | "Sentence_Break"
    | "Sentence_Terminal"
    | "Simple_Case_Folding"
    | "Simple_Lowercase_Mapping"
    | "Simple_Titlecase_Mapping"
    | "Simple_Uppercase_Mapping"
    | "slc"
    | "Soft_Dotted"
    | "stc"
    | "STerm"
    | "suc"
    | "tc"
    | "Term"
    | "Terminal_Punctuation"
    | "Titlecase_Mapping"
    | "uc"
    | "UIdeo"
    | "Unicode_1_Name"
    | "Unified_Ideograph"
    | "Upper"
    | "Uppercase"
    | "Uppercase_Mapping"
    | "Variation_Selector"
    | "Vertical_Orientation"
    | "vo"
    | "VS"
    | "WB"
    | "White_Space"
    | "Word_Break"
    | "WSpace"
    | "XID_Continue"
```

```
    | "XID_Start"
    | "XIDC"
    | "XIDS"
    | "XO_NFC"
    | "XO_NFD"
    | "XO_NFKC"
    | "XO_NFKD"
   },
   attribute value { text },
   empty
}
```

---

**<unihanProp>** (unihan property) holds the name and value of a normative or informative Unihan character (or glyph) property as part of its attributes. [*5.2.1. Character Properties*]

*Module* gaiji — Characters, Glyphs, and Writing Modes

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source

  * *@source*

- att.gaijiProp
  - ~~name~~
  - ~~value~~
  - @version
  - @scheme
- att.datable
  - @period
- att.datable.custom
  - @when-custom
  - @notBefore-custom
  - @notAfter-custom
  - @from-custom
  - @to-custom
  - @datingPoint
  - @datingMethod
- att.datable.iso
  - @when-iso
  - @notBefore-iso
  - @notAfter-iso
  - @from-iso
  - @to-iso
- att.datable.w3c
  - @when
  - @notBefore
  - @notAfter
  - @from
  - @to

`@name` specifies the normalized name of a unicode han database (Unihan) property.

*Status* Required

*Datatype* teidata.xmlName

*Legal values are:* **kZVariant**

**kAccountingNumeric**

**kBigFive**

**kCCCII**

**kCNS1986**

**kCNS1992**

**kCangjie**

**kCantonese**

**kCheungBauer**

**kCheungBauerIndex**

**kCihaiT**

**kCompatibilityVariant**

**kCowles**

**kDaeJaweon**
**kDefinition**
**kEACC**
**kFenn**
**kFennIndex**
**kFourCornerCode**
**kFrequency**
**kGB0**
**kGB1**
**kGB3**
**kGB5**
**kGB7**
**kGB8**
**kGSR**
**kGradeLevel**
**kHDZRadBreak**
**kHKGlyph**
**kHKSCS**
**kHanYu**
**kHangul**
**kHanyuPinlu**
**kHanyuPinyin**
**kIBMJapan**
**kIICore**
**kIRGDaeJaweon**
**kIRGDaiKanwaZiten**
**kIRGHanyuDaZidian**
**kIRGKangXi**
**kIRG_GSource**
**kIRG_HSource**
**kIRG_JSource**
**kIRG_KPSource**
**kIRG_KSource**
**kIRG_MSource**
**kIRG_TSource**
**kIRG_USource**
**kIRG_VSource**
**kJIS0213**
**kJa**
**kJapaneseKun**
**kJapaneseOn**
**kJinmeiyoKanji**
**kJis0**
**kJis1**
**kJoyoKanji**
**kKPS0**

**kKPS1**
**kKSC0**
**kKSC1**
**kKangXi**
**kKarlgren**
**kKorean**
**kKoreanEducationHanja**
**kKoreanName**
**kLau**
**kMainlandTelegraph**
**kMandarin**
**kMatthews**
**kMeyerWempe**
**kMorohashi**
**kNelson**
**kOtherNumeric**
**kPhonetic**
**kPrimaryNumeric**
**kPseudoGB1**
**kRSAdobe_Japan1_6**
**kRSJapanese**
**kRSKanWa**
**kRSKangXi**
**kRSKorean**
**kRSUnicode**
**kSBGY**
**kSemanticVariant**
**kSimplifiedVariant**
**kSpecializedSemanticVariant**
**kTGH**
**kTaiwanTelegraph**
**kTang**
**kTotalStrokes**
**kTraditionalVariant**
**kVietnamese**
**kXHC1983**
**kXerox**

`@value` specifies the value of a named Unihan property
*Status* Required
*Datatype* teidata.word

*Containing entity* char glyph
*May contain* Empty element
*Note* A definitive list of current Unihan property names is provided in the Unicode Han Database.
*Example*

```
<unihanProp name="kRSKangXi" value="120.5"
  version="12.1"/>
```

*Content model* <content> <empty/></content>

*Schema Declaration*

```
element unihanProp
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.gaijiProp.attribute.version,
   att.gaijiProp.attribute.scheme,
   att.datable.attribute.period,
   att.datable.custom.attribute.when-custom,
   att.datable.custom.attribute.notBefore-custom,
   att.datable.custom.attribute.notAfter-custom,
   att.datable.custom.attribute.from-custom,
   att.datable.custom.attribute.to-custom,
   att.datable.custom.attribute.datingPoint,
   att.datable.custom.attribute.datingMethod,
   att.datable.iso.attribute.when-iso,
   att.datable.iso.attribute.notBefore-iso,
   att.datable.iso.attribute.notAfter-iso,
   att.datable.iso.attribute.from-iso,
   att.datable.iso.attribute.to-iso,
   att.datable.w3c.attribute.when,
   att.datable.w3c.attribute.notBefore,
   att.datable.w3c.attribute.notAfter,
   att.datable.w3c.attribute.from,
   att.datable.w3c.attribute.to,
   attribute name
   {
      "kZVariant"
    | "kAccountingNumeric"
    | "kBigFive"
    | "kCCCII"
    | "kCNS1986"
    | "kCNS1992"
    | "kCangjie"
    | "kCantonese"
    | "kCheungBauer"
    | "kCheungBauerIndex"
    | "kCihaiT"
    | "kCompatibilityVariant"
    | "kCowles"
    | "kDaeJaweon"
    | "kDefinition"
    | "kEACC"
    | "kFenn"
    | "kFennIndex"
    | "kFourCornerCode"
    | "kFrequency"
```

```
    | "kGB0"
    | "kGB1"
    | "kGB3"
    | "kGB5"
    | "kGB7"
    | "kGB8"
    | "kGSR"
    | "kGradeLevel"
    | "kHDZRadBreak"
    | "kHKGlyph"
    | "kHKSCS"
    | "kHanYu"
    | "kHangul"
    | "kHanyuPinlu"
    | "kHanyuPinyin"
    | "kIBMJapan"
    | "kIICore"
    | "kIRGDaeJaweon"
    | "kIRGDaiKanwaZiten"
    | "kIRGHanyuDaZidian"
    | "kIRGKangXi"
    | "kIRG_GSource"
    | "kIRG_HSource"
    | "kIRG_JSource"
    | "kIRG_KPSource"
    | "kIRG_KSource"
    | "kIRG_MSource"
    | "kIRG_TSource"
    | "kIRG_USource"
    | "kIRG_VSource"
    | "kJIS0213"
    | "kJa"
    | "kJapaneseKun"
    | "kJapaneseOn"
    | "kJinmeiyoKanji"
    | "kJis0"
    | "kJis1"
    | "kJoyoKanji"
    | "kKPS0"
    | "kKPS1"
    | "kKSC0"
    | "kKSC1"
    | "kKangXi"
    | "kKarlgren"
    | "kKorean"
    | "kKoreanEducationHanja"
    | "kKoreanName"
    | "kLau"
    | "kMainlandTelegraph"
    | "kMandarin"
    | "kMatthews"
    | "kMeyerWempe"
    | "kMorohashi"
    | "kNelson"
    | "kOtherNumeric"
    | "kPhonetic"
    | "kPrimaryNumeric"
    | "kPseudoGB1"
    | "kRSAdobe_Japan1_6"
```

```
    | "kRSJapanese"
    | "kRSKanWa"
    | "kRSKangXi"
    | "kRSKorean"
    | "kRSUnicode"
    | "kSBGY"
    | "kSemanticVariant"
    | "kSimplifiedVariant"
    | "kSpecializedSemanticVariant"
    | "kTGH"
    | "kTaiwanTelegraph"
    | "kTang"
    | "kTotalStrokes"
    | "kTraditionalVariant"
    | "kVietnamese"
    | "kXHC1983"
    | "kXerox"
   },
   attribute value { text },
   empty
}
```

---

**<unit>** contains a symbol, a word or a phrase referring to a unit of measurement in any kind of formal or informal system. [*3.6.3. Numbers and Measures*]

*Module* core — Elements Available in All TEI Documents

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*

  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.cmc
  - *@generatedBy*
- att.measurement
  - *@unit*
  - *@unitRef*
  - *@quantity*
  - *@commodity*
- att.typed
  - *@type*
  - *@subtype*

*Member of* model.measureLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure measureGrp meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage unitDef

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example* Here is an example of a <unit> element holding a *unitRef* attribute that points to a definition of the unit in the TEI header.

```
<measure>
  <num>3</num>
  <unit  unitRef="#ell">ells</unit>
</measure>
<!-- In the TEI Header: -->
<encodingDesc>
  <unitDecl>
    <unitDef  xml:id="ell">
      <label>ell</label>
      <placeName  ref="#iceland"/>
      <desc>A unit of measure for cloth, roughly equivalent to 18 inches, or
from an adult male’s elbow to the tip of the middle finger.</desc>
    </unitDef>
  </unitDecl>
</encodingDesc>
```

*Example*

```
<measure>
  <num>2</num>
  <unit>kg</unit>
</measure>
```

*Example*

```
<measure  type="value">
  <num>3</num>
  <unit  type="time"  unit="min">minute</unit>
</measure>
```

*Example*

```
<measure  type="interval">
  <num  atLeast="1.2">1.2</num> to <num  atMost="5.6">5.6</num>
  <unit  type="velocity"  unit="km/h">km/h</unit>
</measure>
```

*Example*

```
<p>Light travels at <num  value="3E10">3×10^10</num>
  <unit  type="rate"  unit="cm/s">
    <unit  type="space">cm</unit> per <unit  type="time">second</unit>
  </unit>.</p>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element unit
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.measurement.attributes,
   att.typed.attributes,
   macro.phraseSeq}
```

---

**<unitDecl>** (unit declarations) provides information about units of measurement that are not members of the International System of Units. [*2.3.9. The Unit Declaration*]

*Module* header — The TEI Header
*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.canonical
  - *@key*
  - *@ref*
- att.datable
  - *@period*

- att.datable.custom
  - *@when-custom*
  - *@notBefore-custom*
  - *@notAfter-custom*
  - *@from-custom*
  - *@to-custom*
  - *@datingPoint*

  - *@datingMethod*
- att.datable.iso
  - *@when-iso*
  - *@notBefore-iso*
  - *@notAfter-iso*
  - *@from-iso*
  - *@to-iso*
- att.datable.w3c
  - *@when*
  - *@notBefore*
  - *@notAfter*
  - *@from*
  - *@to*

*Member of* model.encodingDescPart
*Contained by*
*header:* encodingDesc
*May contain*
*header:* unitDef
*Example*

```
<unitDecl>
  <unitDef  xml:id="□"  type="weight">
    <label  xml:lang="ja">斤</label>
    <label  xml:lang="ja-Latn">kin</label>
    <conversion  fromUnit="#□"  toUnit="#□"
      formula="16"/>
  </unitDef>
  <unitDef  xml:id="□"  type="weight">
    <label  xml:lang="ja">両</label>
    <label  xml:lang="ja-Latn">ryo</label>
    <conversion  fromUnit="#□"  toUnit="#□"
      formula="4"/>
  </unitDef>
  <unitDef  xml:id="□"  type="weight">
    <label  xml:lang="ja">分</label>
    <label  xml:lang="ja-Latn">Bu</label>
    <conversion  fromUnit="#□"  toUnit="#□"
      formula="6"/>
  </unitDef>
  <unitDef  xml:id="□"  type="weight">
    <label  xml:lang="ja">銖</label>
    <label  xml:lang="ja-Latn">Shu</label>
  </unitDef>
</unitDecl>
```

*Content model*

```
<content>
 <elementRef key="unitDef" minOccurs="1"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element unitDecl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   unitDef+
}
```

---

**<unitDef>** (unit definition) contains descriptive information related to a specific unit of measurement. [*2.3.9. The Unit Declaration*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*

    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.canonical
  - *@key*
  - *@ref*
- att.datable
  - *@period*

  - att.datable.custom
    * *@when-custom*
    * *@notBefore-custom*
    * *@notAfter-custom*
    * *@from-custom*
    * *@to-custom*
    * *@datingPoint*
    * *@datingMethod*
  - att.datable.iso
    * *@when-iso*
    * *@notBefore-iso*
    * *@notAfter-iso*
    * *@from-iso*
    * *@to-iso*
  - att.datable.w3c
    * *@when*
    * *@notBefore*
    * *@notAfter*
    * *@from*
    * *@to*
- att.typed
  - *@type*
  - *@subtype*

*Contained by:* unitDecl
*May contain*
*core:* desc label unit
*header:* conversion
*namesdates:* bloc country district geogName placeName region settlement
*Example*

```
<unitDecl>
  <unitDef  xml:id="pechys"  type="length">
    <label>πῆχυς</label>
    <placeName  ref="#athens"/>
    <conversion  fromUnit="#daktylos"
      toUnit="#pechys"  formula="$fromUnit div 24"/>
    <desc>Equivalent to a cubit or 24 daktyloi.</desc>
  </unitDef>
  <unitDef  xml:id="daktylos"  type="length">
    <label>δάκτυλος</label>
    <placeName  ref="#athens"/>
    <desc>A basic unit of length equivalent to one finger (or the size of a
thumb) in ancient Greece.</desc>
  </unitDef>
</unitDecl>
```

*Content model*

```
<content>
 <alternate minOccurs="1"
  maxOccurs="unbounded">
  <classRef key="model.labelLike"
   minOccurs="1"/>
  <classRef key="model.placeNamePart"
   minOccurs="0"/>
  <elementRef key="conversion"
   minOccurs="0"/>
  <elementRef key="unit" minOccurs="0"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element unitDef
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.canonical.attributes,
   att.datable.attributes,
   att.datable.custom.attributes,
   att.datable.iso.attributes,
   att.datable.w3c.attributes,
   att.typed.attributes,
   ( model.labelLike | model.placeNamePart? | conversion? | unit? )+
}
```

---

**<usg>** (usage) contains usage information in a dictionary entry. [*10.3.5.2. Usage Information and Other Labels*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*
  - *@mergedIn*
  - *@opt*

  - att.datcat
    - *@datcat*
    - *@valueDatcat*
    - *@targetDatcat*

- att.lexicographic.normalized
  - *@norm*
  - *@orig*
- att.typed
  - ~~type~~
  - @subtype

`@type` classifies the usage information using any convenient typology.

*Derived from* att.typed

*Status* Optional

*Datatype* teidata.enumerated

*Sample values include:* **geo** (geographic) geographic area

**time** temporal, historical era (archaic, old, etc.)

**dom** (domain) domain or subject matter (e.g. scientific, literary etc.)

**register** (register)

**style** style (figurative, literal, etc.)

**plev** (preference level) preference level (chiefly, usually, etc.)

**lang** (language) name of a language mentioned in etymological or other linguistic discussion.

**gram** (grammatical) grammatical usage

**syn** (synonym) synonym given to show use

**hyper** (hypernym) hypernym given to show usage

**colloc** (collocation) typical collocate of the headword.

**comp** (complement) typical complement

**obj** (object) typical object

**subj** (subject) typical subject

**verb** typical verb

**hint** unclassifiable piece of information to guide sense choice

*Member of* model.entryPart model.entryPart.top model.lexicalRefinement

*Contained by*

*core:* cit

*dictionaries:* dictScrap entry entryFree etym form gramGrp hom re sense xr

*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<form>
  <orth>colour</orth>
  <form>
    <usg  type="geo">U.S.</usg>
    <orth>color</orth>
  </form>
</form>
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element usg
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   macro.paraContent}
```

---

**<val>** (value) contains a single attribute value. [*23. Documentation Elements 23.5.3. Attribute List Specification*]

*Module* tagdocs — Documentation Elements

*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  - * *@ana*
- – att.global.change
  - * *@change*
- – att.global.facs
  - * *@facs*
- – att.global.linking
  - * *@corresp*
  - * *@synch*
  - * *@sameAs*
  - * *@copyOf*
  - * *@next*
  - * *@prev*
  - * *@exclude*
  - * *@select*
- – att.global.rendition
  - * *@rend*
  - * *@style*
  - * *@rendition*
- – att.global.responsibility
  - * *@cert*
  - * *@resp*
- – att.global.source
  - * *@source*

*Member of* model.phrase.xml

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain* Character data only

*Example*

```
<val>unknown</val>
```

*Content model* `<content> <textNode/></content>`

*Schema Declaration*

```
element val
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   text
}
```

---

## <valDesc>

(value description) specifies any semantic or syntactic constraint on the value that an attribute may take, additional to the information carried by the <datatype> element. [*23.5.3. Attribute List Specification*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.combinable
  - *@mode*

- att.deprecated
  - *@validUntil*
- att.translatable
  - *@versionDate*

*Contained by:* tagdocs: attDef
*May contain*
*analysis:* interp interpGrp span spanGrp
*certainty:* certainty precision respons
*core:* abbr address cb choice date distinct ellipsis email emph expan foreign gap gb gloss hi index lb measure measureGrp mentioned milestone name note noteGrp num pb ptr q ref rs soCalled term time title unit
*dictionaries:* lang
*figures:* figure notatedMusic
*header:* idno
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef tag val

*textcrit:* app witDetail

*transcr:* addSpan am damageSpan delSpan ex fw listTranspose metamark space subst substJoin

character data

*Example*

```
<valDesc>must point to another <gi>align</gi>
element logically preceding this one.</valDesc>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq.limited"/>
</content>
```

*Schema Declaration*

```
element valDesc
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   att.translatable.attributes,
   macro.phraseSeq.limited}
```

---

**<valItem>** documents a single value in a predefined list of values. [*23.5.3. Attribute List Specification*]

*Module* tagdocs — Documentation Elements

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic

* *@ana*

- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.combinable
  - *@mode*

- att.deprecated
  - *@validUntil*

`@ident` specifies the value concerned.
*Status* Required
*Datatype* teidata.text

*Contained by* *tagdocs:* valList
*May contain*
*core:* desc gloss
*tagdocs:* altIdent equiv paramList remarks
*Example*

```
<valItem ident="dub">
  <altIdent xml:lang="fr">dou</altIdent>
  <equiv name="unknown"/>
  <gloss>dubious</gloss>
  <desc>used when the application of this element is doubtful or
uncertain</desc>
</valItem>
```

*Example* The <remarks> element is only allowed within <valItem> if the <desc> element is also present:

```
<valList  type="closed">
  <valItem  ident="border">
    <desc>used to signify that there is border around a page.</desc>
  </valItem>
  <valItem  ident="border-rule">
    <desc>used to signify that there is a border made of ruled lines.</desc>
  </valItem>
  <valItem  ident="border-ornamental">
    <desc>used to signify that there is an ornamental border.</desc>
  </valItem>
  <valItem  ident="catch">
    <gloss>catchword</gloss>
    <desc>for catch words.</desc>
  </valItem>
  <valItem  ident="lineNum">
    <gloss>line number</gloss>
    <desc>for printed line numbers (usually in poetry).</desc>
  </valItem>
  <valItem  ident="listHead">
    <gloss>list heading</gloss>
    <desc>for headings of lists or similar constructs; often used in tables of
contents for the
    <q>page</q> heading.</desc>
  </valItem>
  <valItem  ident="pageNum">
    <gloss>page number</gloss>
    <desc>for page numbers.</desc>
  </valItem>
  <valItem  ident="pressFig">
    <gloss>press figure</gloss>
    <desc>for press figures. These are usually numbers, but are not sequential
or regularly
        printed like signatures.</desc>
  </valItem>
  <valItem  ident="sig">
    <gloss>signature</gloss>
    <desc>for printed signatures.</desc>
  </valItem>
  <valItem  ident="unknown">
    <desc>for a metawork whose function you do not know.</desc>
  </valItem>
  <valItem  ident="vol">
    <gloss>volume number</gloss>
    <desc>for printed volume numbers.</desc>
  </valItem>
  <valItem  ident="other">
    <desc>anything not covered by the above values.</desc>
    <remarks>
      <p>Used for an ornament, decorated tool line, or symbol(s) that is not
officially a
           delimiter (and thus there is no element on which to put
<val>pre()</val> or
     <val>post()</val>) but is used to fill space at the bottom of a page so
that a
     <gi>floatingText</gi> or similar can start at the top of the next.</p>
    </remarks>
```

```
      </valItem>
    </valList>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.identSynonyms"
   minOccurs="0" maxOccurs="unbounded"/>
  <sequence minOccurs="0">
   <classRef key="model.descLike"
    minOccurs="1" maxOccurs="unbounded"/>
   <sequence minOccurs="0"
    maxOccurs="unbounded">
    <classRef key="model.identSynonyms"
     minOccurs="1" maxOccurs="1"/>
    <classRef key="model.descLike"
     minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
   <elementRef key="remarks" minOccurs="0"
    maxOccurs="unbounded"/>
  </sequence>
  <elementRef key="paramList" minOccurs="0"
   maxOccurs="1"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element valItem
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   attribute ident { text },
   (
      model.identSynonyms*,
      (
         (
            model.descLike+,
            ( ( model.identSynonyms, model.descLike* )* ),
            remarks*
         )?
      ),
      paramList?
   )
}
```

**<valList>** (value list) contains one or more <valItem> elements defining possible values. [*23.5.3. Attribute List Specification*]

*Module* tagdocs — Documentation Elements

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.combinable
  - *@mode*
  - att.deprecated
    - *@validUntil*

`@type` specifies the extensibility of the list of values specified.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **closed** only the values specified are permitted.

**semi** (semi-open) all the values specified should be supported, but other values are legal and software should have appropriate fallback processing for them.

**open** the values specified are sample values only.[Default]

*Contained by tagdocs:* alternate attDef content dataSpec elementSpec macroSpec
*May contain*
*tagdocs:* valItem
*Example*

```
<valList type="closed">
  <valItem ident="req">
    <gloss>required</gloss>
  </valItem>
  <valItem ident="rec">
    <gloss>recommended</gloss>
  </valItem>
  <valItem ident="opt">
    <gloss>optional</gloss>
  </valItem>
</valList>
```

*Content model*

```
<content>
 <elementRef key="valItem" minOccurs="0"
  maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element valList
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.combinable.attributes,
   att.deprecated.attributes,
   attribute type { "closed" | "semi" | "open" }?,
   valItem*
}
```

---

## <vAlt>

(value alternation) represents the value part of a feature-value specification which contains a set of values, only one of which can be valid. [*19.8.1. Alternation*]

*Module* iso-fs — Feature Structures
*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*

- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

*Member of* model.featureVal.single
*Contained by*
*iso-fs:* f fvLib if vAlt vColl vDefault vLabel vMerge vNot vRange
*May contain*
*iso-fs:* binary default fs numeric string symbol vAlt vColl vLabel vMerge vNot
*Example*

```
<f  name="gender">
  <vAlt>
    <symbol  value="masculine"/>
    <symbol  value="neuter"/>
    <symbol  value="feminine"/>
  </vAlt>
</f>
```

*Content model*

```
<content>
 <sequence>
  <classRef key="model.featureVal"/>
```

```
  <classRef key="model.featureVal"
   minOccurs="1" maxOccurs="unbounded"/>
 </sequence>
</content>
```

*Schema Declaration*

```
element vAlt
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.featureVal, model.featureVal+ )
}
```

---

## <variantEncoding>

<variantEncoding> (variant encoding) declares the method used to encode text-critical variants. [*13.1.1. The Apparatus Entry*]

*Module* textcrit — Critical Apparatus

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition

* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@method` indicates which method is used to encode the apparatus of variants.
*Status* Required
*Datatype* teidata.enumerated
*Legal values are:* **location-referenced** apparatus uses line numbers or other canonical reference scheme referenced in a base text.
**double-end-point** apparatus indicates the precise locations of the beginning and ending of each lemma relative to a base text.
**parallel-segmentation** alternate readings of a passage are given in parallel in the text; no notion of a base text is necessary.
*Note* The value 'parallel-segmentation' requires in-line encoding of the apparatus.

`@location` indicates whether the apparatus appears within the running text or external to it.
*Status* Required
*Datatype* teidata.enumerated
*Schematron* <sch:rule context="tei:variantEncoding"> <sch:report test="@location eq 'external' and @method eq 'parallel-segmentation'"> The @location value "external" is inconsistent with the parallel-segmentation method of apparatus markup.</sch:report> </sch:rule>
*Legal values are:* **internal** apparatus appears within the running text.
**external** apparatus appears outside the base text.
*Note* The value 'external' is inconsistent with the parallel-segmentation method of apparatus markup.

*Member of* model.encodingDescPart
*Contained by*
*header:* encodingDesc
*May contain* Empty element
*Example*

```
<variantEncoding method="location-referenced"
  location="external"/>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element variantEncoding
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
```

```
   att.global.source.attributes,
   attribute method
   {
      "location-referenced" | "double-end-point" | "parallel-segmentation"
   },
   attribute location { "internal" | "external" },
   empty
}
```

---

**<vColl>** (collection of values) represents the value part of a feature-value specification which contains multiple values organized as a set, bag, or list. [*19.7. Collections as Complex Feature Values*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*

`@org` (organization) indicates organization of given value or values as set, bag or list.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **set** indicates that the given values are organized as a set.

**bag** indicates that the given values are organized as a bag (multiset).

**list** indicates that the given values are organized as a list.

*Member of* model.featureVal.complex

*Contained by*

*iso-fs:* f fvLib if vAlt vColl vDefault vLabel vMerge vNot vRange

*May contain*

*iso-fs:* binary default fs numeric string symbol vAlt vColl vLabel

*Example*

```
<f name="name">
  <vColl>
    <string>Jean</string>
    <string>Luc</string>
    <string>Godard</string>
  </vColl>
</f>
```

*Example*

```
<fs>
  <f name="lex">
    <symbol value="auxquels"/>
  </f>
  <f name="maf">
    <vColl org="list">
      <fs>
        <f name="cat">
          <symbol value="prep"/>
        </f>
      </fs>
      <fs>
        <f name="cat">
          <symbol value="pronoun"/>
        </f>
        <f name="kind">
          <symbol value="rel"/>
        </f>
        <f name="num">
          <symbol value="pl"/>
        </f>
        <f name="gender">
          <symbol value="masc"/>
        </f>
      </fs>
    </vColl>
  </f>
</fs>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
```

```
  maxOccurs="unbounded">
  <elementRef key="fs"/>
  <classRef key="model.featureVal.single"/>
  <elementRef key="vColl"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element vColl
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute org { "set" | "bag" | "list" }?,
   ( fs | model.featureVal.single | vColl )*
}
```

---

**<vDefault>** (value default) declares the default value to be supplied when a feature structure does not contain an instance of <f> for this name; if unconditional, it is specified as one (or, depending on the value of the *org* attribute of the enclosing <fDecl>) more <fs> elements or primitive values; if conditional, it is specified as one or more <if> elements; if no default is specified, or no condition matches, the value none is assumed. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*

* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

*Contained by:* *iso-fs:* fDecl
*May contain*
*iso-fs:* binary default fs if numeric string symbol vAlt vColl vLabel vMerge vNot
*Note* May contain a legal feature value, or a series of <if> elements.
*Example*

```
<fDecl  name="INV">
  <fDescr>inverted sentence</fDescr>
  <vRange>
    <vAlt>
      <binary  value="true"/>
      <binary  value="false"/>
    </vAlt>
  </vRange>
  <vDefault>
    <binary  value="false"/>
  </vDefault>
</fDecl>
```

*Content model*

```
<content>
 <alternate>
  <classRef key="model.featureVal"
   minOccurs="1" maxOccurs="unbounded"/>
  <elementRef key="if" minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element vDefault
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
```

```
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   ( model.featureVal+ | if+ )
}
```

---

**<view>** (view) describes the visual context of some part of a screen play in terms of what the spectator sees, generally independent of any dialogue. [*7.3.1. Technical Information 7.3. Other Types of Performance Text*]

*Module* drama — Performance Texts

*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  - * *@ana*
- – att.global.change
  - * *@change*
- – att.global.facs
  - * *@facs*
- – att.global.linking
  - * *@corresp*
  - * *@synch*
  - * *@sameAs*
  - * *@copyOf*
  - * *@next*
  - * *@prev*
  - * *@exclude*
  - * *@select*
- – att.global.rendition
  - * *@rend*
  - * *@style*
  - * *@rendition*
- – att.global.responsibility
  - * *@cert*
  - * *@resp*
- – att.global.source
  - * *@source*

*Member of* model.stageLike

*Contained by*

*cmc:* post

*core:* add corr del desc emph head hi item l lg meeting note orig p q quote ref reg said sic sp stage textLang title unclear

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron stress subc syll tns usg xr

*drama:* camera caption castList epilogue performance prologue set sound spGrp tech view

*figures:* cell figDesc figure

*header:* change handNote licence rendition scriptNote tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions collation condition custEvent decoNote filiation foliation layout musicNotation origin provenance signatures source summary support surrogates typeNote

*namesdates:* occupation

*spoken:* u writing

*textcrit:* lem rdg witness

*textstructure:* argument body div div1 div2 div3 div4 div5 div6 div7 docEdition epigraph imprimatur postscript salute signed titlePart trailer

*transcr:* damage metamark mod restore retrace secl supplied surplus

*verse:* rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig p pb ptr q quote ref reg rs ruby said sic soCalled sp stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound spGrp tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* ab alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*nets:* eTree forest graph listForest tree

*spoken:* annotationBlock incident kinesic pause shift u vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme
character data

*Note* A view is a particular form of stage direction.

*Example*

```
<view>
  <name>Max</name> joins his daughter
at the window. <hi>Rain</hi> sprays his
face--
</view>
<view>
  <camera>Max's POV</camera> He sees occasional
windows open, and just across from his apartment
house, a <hi>man</hi> opens the front door of
a brownstone--
</view>
```

*Example*

```
<div  type="shot">
  <view>BBC World symbol</view>
  <sp>
    <speaker>Voice Over</speaker>
    <p>Monty Python's Flying Circus tonight comes to you live
        from the Grillomat Snack Bar, Paignton.</p>
  </sp>
</div>
<div  type="shot">
  <view>Interior of a nasty snack bar. Customers around, preferably
     real people. Linkman sitting at one of the plastic tables.</view>
  <sp>
    <speaker>Linkman</speaker>
    <p>Hello to you live from the Grillomat Snack Bar.
    </p>
  </sp>
</div>
```

*Content model*

```
<content>
 <macroRef key="macro.specialPara"/>
</content>
```

*Schema Declaration*

```
element view
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.specialPara}
```

**<vLabel>** (value label) represents the value part of a feature-value specification which appears at more than one point in a feature structure. [*19.6. Re-entrant Feature Structures*]

*Module* iso-fs — Feature Structures

*Attributes* • att.global

- *@xml:id*
- *@n*
- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@name` supplies a name identifying the sharing point.

*Status* Required

*Datatype* teidata.word

*Member of* model.featureVal.single

*Contained by*

*iso-fs:* f fvLib if vAlt vColl vDefault vLabel vMerge vNot vRange

*May contain*

*iso-fs:* binary default fs numeric string symbol vAlt vColl vLabel vMerge vNot

*Example*

```
<fs>
  <f  name="nominal">
    <fs>
      <f  name="nm-num">
        <vLabel  name="L1">
          <symbol  value="singular"/>
        </vLabel>
      </f>
<!-- other nominal features -->
    </fs>
  </f>
  <f  name="verbal">
    <fs>
      <f  name="vb-num">
        <vLabel  name="L1"/>
      </f>
    </fs>
<!-- other verbal features -->
  </f>
</fs>
```

*Content model*

```
<content>
 <classRef key="model.featureVal"
  minOccurs="0"/>
</content>
```

*Schema Declaration*

```
element vLabel
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute name { text },
   model.featureVal?
}
```

---

**<vMerge>** (merged collection of values) represents a feature value which is the result of merging together the feature values contained by its children, using the organization specified by the *org* attribute. [*19.8.3. Collection of Values*]

*Module* iso-fs — Feature Structures

*Attributes*
- att.global
  - *@xml:id*
  - *@n*

- *@xml:lang*
- *@xml:base*
- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

`@org` indicates the organization of the resulting merged values as set, bag or list.

*Status* Optional

*Datatype* teidata.enumerated

*Legal values are:* **set** indicates that the resulting values are organized as a set.

**bag** indicates that the resulting values are organized as a bag (multiset).

**list** indicates that the resulting values are organized as a list.

*Member of* model.featureVal.complex

*Contained by*

*iso-fs:* f fvLib if vAlt vDefault vLabel vMerge vNot vRange

*May contain*

*iso-fs:* binary default fs numeric string symbol vAlt vColl vLabel vMerge vNot

*Example*

```
<vMerge org="list">
  <vColl org="set">
    <symbol value="masculine"/>
```

```
      <symbol value="neuter"/>
      <symbol value="feminine"/>
    </vColl>
    <symbol value="indeterminate"/>
  </vMerge>
```

This example returns a list, concatenating the indeterminate value with the set of values masculine, neuter and feminine.

*Content model*

```
<content>
 <classRef key="model.featureVal"
  minOccurs="1" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element vMerge
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute org { "set" | "bag" | "list" }?,
   model.featureVal+
}
```

---

**<vNot>** (value negation) represents a feature value which is the negation of its content. [*19.8.2. Negation*]

*Module* iso-fs — Feature Structures

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*

*Member of* model.featureVal.complex

*Contained by*

*iso-fs:* f fvLib if vAlt vDefault vLabel vMerge vNot vRange

*May contain*

*iso-fs:* binary default fs numeric string symbol vAlt vColl vLabel vMerge vNot

*Example*

```
<vNot>
  <symbol  value="masculine"/>
</vNot>
```

*Example*

```
<f  name="mode">
  <vNot>
    <vAlt>
      <symbol  value="infinitive"/>
      <symbol  value="participle"/>
    </vAlt>
  </vNot>
</f>
```

*Content model*

```
<content>
 <classRef key="model.featureVal"/>
</content>
```

*Schema Declaration*

```
element vNot
{
   att.global.attributes,
   att.global.analytic.attributes,
```

```
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   model.featureVal}
```

---

**<vocal>** (vocal) marks any vocalized but not necessarily lexical phenomenon, for example voiced pauses, non-lexical backchannels, etc. [*8.3.3. Vocal, Kinesic, Incident*]

*Module* spoken — Transcriptions of Speech

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.ascribed.directed

- *@toWhom*

- att.ascribed
  - *@who*
- att.cmc
  - *@generatedBy*
- att.timed
  - *@start*
  - *@end*
- att.duration
  - att.duration.iso
    - *@dur-iso*
  - att.duration.w3c
    - *@dur*
- att.typed
  - *@type*
  - *@subtype*

`@iterated` indicates whether or not the phenomenon is repeated.

*Status* Optional

*Datatype* teidata.xTruthValue

*Default* false

*Note* The value true indicates that the vocal effect is repeated several times rather than just occurring once.

*Member of* model.global.spoken

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName

occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* annotationBlock u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*core:* desc

*Example*

```
<vocal dur="PT12S">
  <desc>whistles</desc>
</vocal>
<vocal iterated="true">
  <desc>whistles intermittently</desc>
</vocal>
```

*Content model*

```
<content>
 <classRef key="model.descLike"
  minOccurs="0" maxOccurs="unbounded"/>
</content>
```

*Schema Declaration*

```
element vocal
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.directed.attributes,
   att.ascribed.attributes,
   att.cmc.attributes,
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   att.typed.attributes,
   attribute iterated { text }?,
   model.descLike*
}
```

**<vRange>** (value range) defines the range of allowed values for a feature, in the form of an <fs>, <vAlt>, or primitive value; for the value of an <f> to be valid, it must be subsumed by the specified range; if the <f> contains multiple values (as sanctioned by the *org* attribute), then each value must be subsumed by the <vRange>. [*19.11. Feature System Declaration*]

*Module* iso-fs — Feature Structures

*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  - * *@ana*
- – att.global.change
  - * *@change*
- – att.global.facs
  - * *@facs*
- – att.global.linking
  - * *@corresp*
  - * *@synch*
  - * *@sameAs*
  - * *@copyOf*
  - * *@next*
  - * *@prev*
  - * *@exclude*
  - * *@select*
- – att.global.rendition
  - * *@rend*
  - * *@style*
  - * *@rendition*
- – att.global.responsibility
  - * *@cert*
  - * *@resp*
- – att.global.source
  - * *@source*

*Contained by* *iso-fs:* fDecl

*May contain*

*iso-fs:* binary default fs numeric string symbol vAlt vColl vLabel vMerge vNot

*Note* May contain any legal feature-value specification.

*Example*

```
<fDecl  name="INV">
  <fDescr>inverted sentence</fDescr>
```

```
      <vRange>
        <vAlt>
          <binary value="true"/>
          <binary value="false"/>
        </vAlt>
      </vRange>
      <vDefault>
        <binary value="false"/>
      </vDefault>
    </fDecl>
```

*Content model*

```
<content>
 <classRef key="model.featureVal"/>
</content>
```

*Schema Declaration*

```
element vRange
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   model.featureVal}
```

---

**<W>** (word) represents a grammatical (not necessarily orthographic) word. [*18.1. Linguistic Segment Categories* *18.4.2. Lightweight Linguistic Annotation*]

*Module* analysis — Simple Analytic Mechanisms

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.cmc
  - *@generatedBy*
- att.linguistic
  - *@lemma*
  - *@lemmaRef*
  - *@pos*
  - *@msd*
  - *@join*

  - att.lexicographic.normalized
    * *@norm*
    * *@orig*
- att.notated
  - *@notation*
- att.segLike
  - *@function*

  - att.datcat
    * *@datcat*
    * *@valueDatcat*
    * *@targetDatcat*
  - att.fragmentable
    * *@part*
  - att.metrical
    * *@met*
    * *@real*
    * *@rhyme*

- att.typed
  - *@type*
  - *@subtype*

*Member of* model.linePart model.segLike

*Contained by*

*analysis:* cl phr s w

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure mentioned name note num orig p pubPlace publisher q quote rb ref reg rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell

*header:* change distributor edition extent geoDecl handNote licence scriptNote

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit layout material musicNotation objectType origDate origPlace origin provenance rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg

*textcrit:* lem rdg wit witDetail

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl supplied surplus zone

*verse:* rhyme

*May contain*

*analysis:* c interp interpGrp m pc span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add cb choice corr del ellipsis expan gap gb hi index lb milestone note noteGrp orig pb q reg sic unclear

*figures:* figure notatedMusic

*gaiji:* g

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*spoken:* incident kinesic pause shift vocal writing

*textcrit:* app witDetail

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example* This example is adapted from the Folger Library's Early Modern English Drama version of The Wits: a Comedy by William Davenant.

```
<l>
  <w  lemma="it"  pos="pn"
    xml:id="A19883-003-a-0100">IT</w>
  <w  lemma="have"  pos="vvz"
    xml:id="A19883-003-a-0110">hath</w>
  <w  lemma="be"  pos="vvn"
    xml:id="A19883-003-a-0120">been</w>
  <w  lemma="say"  pos="vvn"
    xml:id="A19883-003-a-0130">said</w>
  <w  lemma="of"  pos="acp-p"
    xml:id="A19883-003-a-0140">of</w>
  <w  lemma="old"  pos="j"
    xml:id="A19883-003-a-0150">old</w>
  <pc  xml:id="A19883-003-a-0160">,</pc>
  <w  lemma="that"  pos="cs"
    xml:id="A19883-003-a-0170">that</w>
  <w  lemma="play"  pos="vvz"
    xml:id="A19883-003-a-0180">
    <choice>
      <orig>Playes</orig>
      <reg>Plays</reg>
    </choice>
  </w>
  <w  lemma="be"  pos="vvb"
    xml:id="A19883-003-a-0190">are</w>
  <w  lemma="feast"  pos="n2"
    xml:id="A19883-003-a-0200">Feasts</w>
  <pc  xml:id="A19883-003-a-0210">,</pc>
</l>
<l  xml:id="A19883-e100220">
  <w  lemma="poet"  pos="n2"
    xml:id="A19883-003-a-0220">Poets</w>
  <w  lemma="the"  pos="d"
    xml:id="A19883-003-a-0230">the</w>
  <w  lemma="cook"  pos="n2"
    xml:id="A19883-003-a-0240">
    <choice>
      <orig>Cookes</orig>
      <reg>Cooks</reg>
    </choice>
  </w>
  <pc  xml:id="A19883-003-a-0250">,</pc>
  <w  lemma="and"  pos="cc"
    xml:id="A19883-003-a-0260">and</w>
  <w  lemma="the"  pos="d"
    xml:id="A19883-003-a-0270">the</w>
  <w  lemma="spectator"  pos="n2"
    xml:id="A19883-003-a-0280">Spectators</w>
  <w  lemma="guest"  pos="n2"
    xml:id="A19883-003-a-0290">Guests</w>
  <pc  xml:id="A19883-003-a-0300">,</pc>
</l>
<l  xml:id="A19883-e100230">
  <w  lemma="the"  pos="d"
    xml:id="A19883-003-a-0310">The</w>
  <w  lemma="actor"  pos="n2"
    xml:id="A19883-003-a-0320">Actors</w>
```

```
      <w  lemma="waiter"  pos="n2"
        xml:id="A19883-003-a-0330">Waiters</w>
      <pc  xml:id="A19883-003-a-0340">:</pc>
    <!-- ... -->
    </l>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <elementRef key="seg"/>
  <elementRef key="w"/>
  <elementRef key="m"/>
  <elementRef key="c"/>
  <elementRef key="pc"/>
  <classRef key="model.global"/>
  <classRef key="model.lPart"/>
  <classRef key="model.hiLike"/>
  <classRef key="model.pPart.edit"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element w
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.cmc.attributes,
   att.linguistic.attributes,
   att.lexicographic.normalized.attributes,
   att.notated.attributes,
   att.segLike.attributes,
   att.datcat.attributes,
   att.fragmentable.attributes,
   att.metrical.attributes,
   att.typed.attributes,
   (
      text
    | model.gLike    | seg    | w    | m    | c    | pc    | model.global    | model.lPart
}
```

---

**<watermark>** (watermark) contains a word or phrase describing a watermark or similar device. [*11.3.3. Watermarks and Stamps*]

*Module* msdescription — Manuscript Description
*Attributes* • att.global
- – *@xml:id*
- – *@n*
- – *@xml:lang*
- – *@xml:base*
- – *@xml:space*

- – att.global.analytic
  * *@ana*
- – att.global.change
  * *@change*
- – att.global.facs
  * *@facs*
- – att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
- – att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
- – att.global.responsibility
  * *@cert*
  * *@resp*
- – att.global.source
  * *@source*

*Member of* model.pPart.msdesc
*Contained by*
*analysis:* cl phr s span
*cmc:* post
*core:* abbr add addrLine author biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit
*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose
*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr
*drama:* actor camera caption castItem role roleDesc sound tech view
*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Example*

```
<support>
  <p>
    <material>Rag paper</material> with <watermark>anchor</watermark>
```

```
watermark</p>
</support>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element watermark
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   macro.phraseSeq}
```

---

**<when>** indicates a point in time either relative to other elements in the same timeline tag, or absolutely. [*17.4.2. Placing Synchronous Events in Time*]

*Module* linking — Linking, Segmentation, and Alignment

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*

* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

`@absolute` supplies an absolute value for the time.
*Status* Optional
*Datatype* teidata.temporal.w3c
*Note* This attribute should always be specified on a <when> element which serves as the target for the *origin* attribute of a <timeline>.

`@unit` specifies the unit of time in which the *interval* value is expressed, if this is not inherited from the parent <timeline>.
*Status* Optional
*Datatype* teidata.enumerated
*Suggested values include:* **d** (days)
**h** (hours)
**min** (minutes)
**s** (seconds)
**ms** (milliseconds)

`@interval` specifies a time interval either as a number or as one of the keywords defined by the datatype `teidata.interval`.
*Status* Optional
*Datatype* teidata.interval

`@since` identifies the reference point for determining the time of the current <when> element, which is obtained by adding the interval to the time of the reference point.
*Status* Optional
*Datatype* teidata.pointer
*Note* This attribute should point to another <when> element in the same <timeline>. If no value is supplied, and the *absolute* attribute is also unspecified, then the reference point is understood to be the origin of the enclosing <timeline> tag.

*Contained by:* timeline
*May contain* Empty element
*Note* On this element, the global *xml:id* attribute must be supplied to specify an identifier for this point in time. The value used may be chosen freely provided that it is unique within the document and is a syntactically valid name. There is no requirement for values containing numbers to be in sequence.
*Example*

```
<when xml:id="TW3" interval="20" since="#w2"/>
```

*Content model* `<content> <empty/></content>`
*Schema Declaration*

```
element when
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   attribute absolute { text }?,
   attribute unit { "d" | "h" | "min" | "s" | "ms" }?,
   attribute interval { text }?,
   attribute since { text }?,
   empty
}
```

---

**<width>** (width) contains a measurement of an object along the axis parallel to its bottom, e.g. perpendicular to the spine of a book or codex. [*11.3.4. Dimensions*]

*Module* msdescription — Manuscript Description

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*

    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.dimensions
  - *@unit*
  - *@quantity*
  - *@extent*
  - *@precision*
  - *@scope*

  - att.ranging
    * *@atLeast*
    * *@atMost*
    * *@min*
    * *@max*
    * *@confidence*

*Member of* model.dimLike model.measureLike

*Contained by*

*analysis:* cl phr s span

*cmc:* post

*core:* abbr add addrLine author bibl biblScope citedRange corr date del desc distinct editor email emph expan foreign gloss head headItem headLabel hi item l label measure measureGrp meeting mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said sic soCalled speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entryFree etym form gen gram gramGrp hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castItem role roleDesc sound tech view

*figures:* cell figDesc

*header:* authority catDesc change classCode creation distributor edition extent funder geoDecl handNote language licence principal rendition scriptNote sponsor tagUsage

*iso-fs:* fDescr fsDescr

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote dimensions explicit filiation finalRubric foliation heraldry incipit institution layout material musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown location nameLink nationality objectName occupation offset orgName persName persPronouns placeName region residence roleName settlement sex socecStatus surname

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail witness

*textstructure:* byline closer dateline docAuthor docDate docEdition docImprint imprimatur opener salute signed titlePart trailer

*transcr:* damage fw metamark mod restore retrace secl supplied surplus

*verse:* metSym rhyme

*May contain*

*gaiji:* g

character data

*Note* If used to specify the depth of a non text-bearing portion of some object, for example a monument, this element conventionally refers to the axis facing the observer, and perpendicular to that indicated by the 'depth' axis.

*Example*

```
<width unit="in">4</width>
```

*Content model* `<content> <macroRef key="macro.xtext"/></content>`

*Schema Declaration*

```
element width
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.dimensions.attributes,
   att.ranging.attributes,
   macro.xtext}
```

---

## <wit>

<wit> (wit) contains a list of one or more sigla of witnesses attesting a given reading, in a textual variation. [*13.1.4. Witness Information*]

*Module* textcrit — Critical Apparatus

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking

    * *@corresp*
    * *@synch*
    * *@sameAs*
    * *@copyOf*
    * *@next*
    * *@prev*
    * *@exclude*
    * *@select*
  - att.global.rendition
    * *@rend*
    * *@style*
    * *@rendition*
  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.rdgPart
  - *@wit*

*Member of* model.rdgPart

*Contained by*

*textcrit:* app lem rdg rdgGrp

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address binaryObject cb choice cit corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* This element represents the same information as that provided by the *wit* attribute of the reading; it may be used to record the exact form of the sigla given in the source edition, when that is of interest.

*Example*

```
<rdg  wit="#El #Hg">Experience</rdg>
<wit>Ellesmere, Hengwryt</wit>
```

*Content model*

```
<content>
 <macroRef key="macro.phraseSeq"/>
</content>
```

*Schema Declaration*

```
element wit
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.rdgPart.attributes,
   macro.phraseSeq}
```

---

**<witDetail>** (witness detail) gives further information about a particular witness, or witnesses, to a particular reading. [*13.1. The Apparatus Entry, Readings, and Witnesses*]

*Module* textcrit — Critical Apparatus

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*

* *@synch*
* *@sameAs*
* *@copyOf*
* *@next*
* *@prev*
* *@exclude*
* *@select*

– att.global.rendition
* *@rend*
* *@style*
* *@rendition*

– att.global.responsibility
* *@cert*
* *@resp*

– att.global.source
* *@source*

• att.placement
– *@place*

• att.pointing
– *@targetLang*
– *@target*
– *@evaluate*

• att.typed
– ~~type~~
– @subtype

`@type` describes the type of information given about the witness.
*Derived from* att.typed
*Status* Optional
*Datatype* teidata.enumerated

`@wit` (witnesses) indicates the sigil or sigla identifying the witness or witnesses to which the detail refers.
*Status* Required
*Datatype* 1–∞ occurrences of teidata.pointer separated by whitespace

*Member of* model.global.edit

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* u writing

*tagdocs:* eg valDesc

*textcrit:* app lem rdg rdgGrp wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl binaryObject cb choice corr date del distinct ellipsis email emph expan foreign gap gb gloss graphic hi index lb measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q ref reg rs ruby sic soCalled term time title unclear unit

*dictionaries:* lang oRef pRef

*figures:* figure formula notatedMusic

*gaiji:* g

*header:* idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att code gi ident listRef specDesc specList tag val

*textcrit:* app witDetail

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The <witDetail> element was formerly permitted anywhere that <note> could appear, but since it should

only be used in association with <lem> and <rdg>, it is recommended that it be placed immediately following the reading that it modifies, in the same <app>. A <witDetail> without a *target* attribute should be assumed to refer to the closest preceding <lem> or <rdg>.

*Example*

```
<app type="substantive">
  <lem xml:id="W026x" wit="#El #HG">Experience</lem>
  <rdg wit="#Ha4">Experiens</rdg>
  <witDetail target="#W026x" resp="#PR"
    wit="#El" type="presentation">Ornamental capital.</witDetail>
</app>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.global"/>
  <elementRef key="bibl"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element witDetail
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.placement.attributes,
   att.pointing.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   attribute wit { list { + } },
   ( text | model.gLike | model.phrase | model.global | bibl )*
}
```

---

**<witEnd>** (fragmented witness end) indicates the end, or suspension, of the text of a fragmentary witness.
[*13.1.5. Fragmentary Witnesses*]

*Module* textcrit — Critical Apparatus

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*

- *@xml:space*

- att.global.analytic
  - *@ana*
- att.global.change
  - *@change*
- att.global.facs
  - *@facs*
- att.global.linking
  - *@corresp*
  - *@synch*
  - *@sameAs*
  - *@copyOf*
  - *@next*
  - *@prev*
  - *@exclude*
  - *@select*
- att.global.rendition
  - *@rend*
  - *@style*
  - *@rendition*
- att.global.responsibility
  - *@cert*
  - *@resp*
- att.global.source
  - *@source*

- att.rdgPart
  - *@wit*

*Member of* model.rdgPart

*Contained by*

*textcrit:* lem rdg

*May contain* Empty element

*Example*

```
<app>
  <lem  wit="#El #Hg">Experience</lem>
  <rdg  wit="#Ha4">Ex<g  ref="#per"/>
    <witEnd/>
  </rdg>
</app>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element witEnd
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.rdgPart.attributes,
   empty
}
```

---

**<witness>** (witness) contains either a description of a single witness referred to within the critical apparatus, or a list of witnesses which is to be referred to by a single sigil. [*13.1. The Apparatus Entry, Readings, and Witnesses*]

*Module* textcrit — Critical Apparatus

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source

  * *@source*

- att.sortable
  - *@sortKey*

*Contained by* *tei:*textcrit listWit

*May contain*

*core:* abbr address bibl biblStruct choice cit date desc distinct email emph expan foreign gloss hi label list listBibl measure measureGrp mentioned name note num ptr q quote ref rs said soCalled stage term time title unit

*dictionaries:* lang

*drama:* camera caption castList move sound tech view

*figures:* table

*header:* biblFull idno

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink object objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specGrp specGrpRef tag val

*textcrit:* listApp listWit

*textstructure:* floatingText

*transcr:* am ex subst

character data

*Note* The content of the <witness> element may give bibliographic information about the witness or witness group, or it may be empty.

*Example*

```
<listWit>
  <witness  xml:id="EL">Ellesmere, Huntingdon Library 26.C.9</witness>
  <witness  xml:id="HG">Hengwrt, National Library of Wales,
     Aberystwyth, Peniarth 392D</witness>
  <witness  xml:id="RA2">Bodleian Library Rawlinson Poetic 149
     (see further
<ptr  target="http://www.examples.com/MSdescs#MSRP149"/>)</witness>
</listWit>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.limitedPhrase"/>
  <classRef key="model.inter"/>
  <elementRef key="note"/>
  <elementRef key="object"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element witness
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.sortable.attributes,
   ( text | model.limitedPhrase | model.inter | note | object )*
}
```

---

**<witStart>** (fragmented witness start) indicates the beginning, or resumption, of the text of a fragmentary witness. [*13.1.5. Fragmentary Witnesses*]

*Module* textcrit — Critical Apparatus

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility

    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.rdgPart
  - *@wit*

*Member of* model.rdgPart

*Contained by*

*textcrit:* lem rdg

*May contain* Empty element

*Example*

```
<app>
  <lem  wit="#El #Hg">Auctoritee</lem>
  <rdg  wit="#La #Ra2">auctorite</rdg>
  <rdg  wit="#X">
    <witStart/>auctorite</rdg>
</app>
```

*Content model* `<content> <empty/></content>`

*Schema Declaration*

```
element witStart
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.rdgPart.attributes,
   empty
}
```

---

## <writing>

<writing> (writing) contains a passage of written text revealed to participants in the course of a spoken text. [*8.3.4. Writing*]

*Module* spoken — Transcriptions of Speech

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    * *@ana*
  - att.global.change

  * *@change*
  - att.global.facs
  * *@facs*
  - att.global.linking
  * *@corresp*
  * *@synch*
  * *@sameAs*
  * *@copyOf*
  * *@next*
  * *@prev*
  * *@exclude*
  * *@select*
  - att.global.rendition
  * *@rend*
  * *@style*
  * *@rendition*
  - att.global.responsibility
  * *@cert*
  * *@resp*
  - att.global.source
  * *@source*
- att.ascribed.directed
  - *@toWhom*

  - att.ascribed
  * *@who*
- att.cmc
  - *@generatedBy*
- att.timed
  - *@start*
  - *@end*

  - att.duration
  * att.duration.iso
  · *@dur-iso*
  * att.duration.w3c
  · *@dur*
- att.typed
  - *@type*
  - *@subtype*

`@gradual` indicates whether the writing is revealed all at once or gradually.
*Status* Optional
*Datatype* teidata.xTruthValue

*Note* The value true indicates the writing is revealed gradually; the value false that the writing is revealed all at once.

*Member of* model.global.spoken

*Contained by*

*analysis:* cl m phr s span w

*cmc:* post

*core:* abbr add addrLine address author bibl biblScope cit citedRange corr date del distinct editor email emph expan foreign gloss head headItem headLabel hi imprint item l label lg list measure mentioned name note num orig p pubPlace publisher q quote rb ref reg resp rs rt said series sic soCalled sp speaker stage street term textLang time title unclear unit

*corpus:* activity channel constitution derivation domain factuality interaction locale preparedness purpose

*dictionaries:* case colloc def dictScrap entry entryFree etym form gen gram gramGrp hom hyph iType lang lbl mood number orth per pos pron re sense stress subc syll tns usg xr

*drama:* actor camera caption castGroup castItem castList epilogue performance prologue role roleDesc set sound spGrp tech view

*figures:* cell figure table

*header:* authority change classCode distributor edition extent funder geoDecl handNote language licence principal scriptNote sponsor

*linking:* ab seg

*msdescription:* accMat acquisition additions catchwords collation collection colophon condition custEvent decoNote explicit filiation finalRubric foliation heraldry incipit institution layout material msItem musicNotation objectType origDate origPlace origin provenance repository rubric secFol signatures source stamp summary support surrogates typeNote watermark

*namesdates:* addName affiliation age birth bloc country death district education eventName faith floruit forename genName gender geogFeat geogName langKnown nameLink nationality objectName occupation offset orgName persName persPronouns person personGrp persona placeName region residence roleName settlement sex socecStatus surname

*nets:* graph

*spoken:* annotationBlock u writing

*tagdocs:* eg valDesc

*textcrit:* lem rdg wit witDetail

*textstructure:* argument back body byline closer dateline div div1 div2 div3 div4 div5 div6 div7 docAuthor docDate docEdition docImprint docTitle epigraph floatingText front group imprimatur opener postscript salute signed text titlePage titlePart trailer

*transcr:* damage fw line metamark mod restore retrace secl sourceDoc supplied surface surfaceGrp surplus zone

*verse:* metSym rhyme

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index l label lb lg list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang oRef pRef

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* The <writing> element will usually be short and most simply transcribed as a character string; the content model also allows a sequence of paragraphs and paragraph-level elements, in case the writing has enough internal structure to warrant such markup. In either case the usual phrase-level tags for written text are available.

*Example*

```
<!-- ... --><l>man in a coonskin cap</l>
<writing>coonskin</writing>
<l>in a pig pen</l>
<writing>pig pen</writing>
<l>wants eleven dollar bills</l>
<writing>20 dollar bills</writing>
<l>you only got ten</l>
<writing>10</writing>
<!-- ... -->
```

*Content model*

```
<content>
 <macroRef key="macro.paraContent"/>
</content>
```

*Schema Declaration*

```
element writing
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.ascribed.directed.attributes,
   att.ascribed.attributes,
   att.cmc.attributes,
```

```
   att.timed.attributes,
   att.duration.iso.attributes,
   att.duration.w3c.attributes,
   att.typed.attributes,
   attribute gradual { text }?,
   macro.paraContent}
```

---

**<xenoData>** (non-TEI metadata) provides a container element into which metadata in non-TEI formats may be placed. [*2.5. Non-TEI Metadata*]

*Module* header — The TEI Header

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.declarable
  - *@default*

- att.typed
  - *@type*
  - *@subtype*

*Member of* model.standOffPart model.teiHeaderPart

*Contained by*

*header:* teiHeader

*linking:* standOff

*May contain* ANY

*Example* This example presumes that the prefix `dc` has been bound to the namespace `http://purl.org/dc/elements/1.1/` and the prefix `rdf` is bound to the namespace `http://www.w3.org/1999/02/22-rdf-syntax-ns#`. Note: The *about* attribute on the `<rdf:Description>` in this example gives a URI indicating the resource to which the metadata contained therein refer. The `<rdf:Description>` in the second <xenoData> block has a blank *about*, meaning it is pointing at the *current* document, so the RDF is about the document within which it is contained, i.e. the TEI document containing the <xenoData> block. Similarly, any kind of relative URI may be used, including fragment identifiers (see *v.6.2 Identifiers and Indicators*). Do note, however, that if the contents of the <xenoData> block are to be extracted and used elsewhere, any relative URIs will have to be resolved accordingly.

```
<xenoData
   xmlns:dc="http://purl.org/dc/elements/1.1/"
   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:RDF>
    <rdf:Description  rdf:about="http://www.worldcat.org/oclc/606621663">
      <dc:title>The description of a new world, called the
blazing-world</dc:title>
     <dc:creator>The Duchess of Newcastle</dc:creator>
     <dc:date>1667</dc:date>
     <dc:identifier>British Library, 8407.h.10</dc:identifier>
     <dc:subject>utopian fiction</dc:subject>
   </rdf:Description>
  </rdf:RDF>
</xenoData>
<xenoData>
  <rdf:RDF>
    <rdf:Description  rdf:about="">
      <dc:title>The Description of a New World, Called the Blazing-World,
1668</dc:title>
     <dc:creator>Cavendish, Margaret (Lucas), Duchess of Newcastle</dc:creator>
     <dc:publisher>Women Writers Project</dc:publisher>
     <dc:date>2002-02-12</dc:date>
     <dc:subject>utopian fiction</dc:subject>
   </rdf:Description>
  </rdf:RDF>
</xenoData>
```

*Example* In this example, the prefix `rdf` is bound to the namespace `http://www.w3.org/1999/02/22-rdf-syntax-ns#`, the prefix `dc` is bound to the namespace `http://purl.org/dc/elements/1.1/`, and the prefix `cc` is bound to the namespace `http://web.resource.org/cc/`.

```
<xenoData
   xmlns:cc="http://web.resource.org/cc/"
   xmlns:dc="http://purl.org/dc/elements/1.1/"
   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:RDF>
```

```
    <cc:Work  rdf:about="">
      <dc:title>Applied Software Project Management - review</dc:title>
      <dc:type  rdf:resource="http://purl.org/dc/dcmitype/Text"/>
      <dc:license  rdf:resource="http://creativecommons.org/licenses/by-sa/2.0/uk/"/>
    </cc:Work>
    <cc:License  rdf:about="http://creativecommons.org/licenses/by-sa/2.0/uk/">
      <cc:permits  rdf:resource="http://web.resource.org/cc/Reproduction"/>
      <cc:permits  rdf:resource="http://web.resource.org/cc/Distribution"/>
      <cc:requires  rdf:resource="http://web.resource.org/cc/Notice"/>
      <cc:requires  rdf:resource="http://web.resource.org/cc/Attribution"/>
      <cc:permits  rdf:resource="http://web.resource.org/cc/DerivativeWorks"/>
      <cc:requires  rdf:resource="http://web.resource.org/cc/ShareAlike"/>
    </cc:License>
  </rdf:RDF>
</xenoData>
```

*Example* In this example, the prefix `dc` is again bound to the namespace `http://www.openarchives.org/OAI/2.0/oai_dc/`, and the prefix `oai_dc` is bound to the namespace `http://www.openarchives.org/OAI/2.0/oai_dc/`.

```
<xenoData
   xmlns:dc="http://purl.org/dc/elements/1.1/"
   xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/">
  <oai_dc:dc>
    <dc:title>The colonial despatches of Vancouver Island and British
        Columbia 1846-1871: 11566, CO 60/2, p. 291; received 13 November.
        Trevelyan to Merivale (Permanent Under-Secretary)</dc:title>
    <dc:date>1858-11-12</dc:date>
    <dc:creator>Trevelyan</dc:creator>
    <dc:publisher>University of Victoria Humanities Computing and Media
        Centre, and UVic Libraries</dc:publisher>
    <dc:type>InteractiveResource</dc:type>
    <dc:format>application/xhtml+xml</dc:format>
    <dc:type>text</dc:type>
    <dc:identifier>http://bcgenesis.uvic.ca/getDoc.htm?id=B585TE13.scx</dc:identifier>
    <dc:rights>This document is licensed under a Creative Commons …</dc:rights>
    <dc:language>(SCHEME=ISO639) en</dc:language>
    <dc:source>Transcribed from microfilm and/or original documents, and
        marked up in TEI P5 XML. The interactive XHTML resource is generated
        from the XHTML using XQuery and XSLT.</dc:source>
    <dc:source>repository: CO</dc:source>
    <dc:source>coNumber: 60</dc:source>
    <dc:source>coVol: 2</dc:source>
    <dc:source>page: 291</dc:source>
    <dc:source>coRegistration: 11566</dc:source>
    <dc:source>received: received 13 November</dc:source>
    <dc:subject>Trevelyan, Sir Charles Edward</dc:subject>
    <dc:subject>Merivale, Herman</dc:subject>
    <dc:subject>Elliot, T. Frederick</dc:subject>
    <dc:subject>Moody, Colonel Richard Clement</dc:subject>
    <dc:subject>Lytton, Sir Edward George Earle Bulwer</dc:subject>
    <dc:subject>Jadis, Vane</dc:subject>
    <dc:subject>Carnarvon, Earl</dc:subject>
    <dc:subject>British Columbia</dc:subject>
    <dc:description>British Columbia correspondence: Public Offices
        document (normally correspondence between government
        departments)</dc:description>
```

```
  </oai_dc:dc>
</xenoData>
```

*Example* In this example, the prefix `mods` is bound to the namespace `http://www.loc.gov/mods/v3`.

```
<xenoData
   xmlns:mods="http://www.loc.gov/mods/v3">
  <mods:mods>
    <mods:titleInfo>
      <mods:title>Academic adaptation and cross-cultural
            learning experiences of Chinese students at American
            universities</mods:title>
      <mods:subTitle>a narrative inquiry</mods:subTitle>
    </mods:titleInfo>
    <mods:name  type="personal"
      authority="local">
      <mods:namePart/>
      <mods:role>
        <mods:roleTerm  authority="marcrelator"
          type="text">Author</mods:roleTerm>
      </mods:role>
      <mods:affiliation>Northeastern University</mods:affiliation>
      <mods:namePart  type="given">Hong</mods:namePart>
      <mods:namePart  type="family">Zhang</mods:namePart>
    </mods:name>
    <mods:name  type="personal"
      authority="local">
      <mods:namePart/>
      <mods:role>
        <mods:roleTerm  authority="local"
          type="text">Advisor</mods:roleTerm>
      </mods:role>
      <mods:namePart  type="given">Liliana</mods:namePart>
      <mods:namePart  type="family">Meneses</mods:namePart>
    </mods:name>

<!-- ... -->
    <mods:typeOfResource>text</mods:typeOfResource>
    <mods:genre>doctoral theses</mods:genre>
    <mods:originInfo>
      <mods:place>
        <mods:placeTerm  type="text">Boston (Mass.)</mods:placeTerm>
      </mods:place>
      <mods:publisher>Northeastern University</mods:publisher>
      <mods:copyrightDate  encoding="w3cdtf"
        keyDate="yes">2013</mods:copyrightDate>
    </mods:originInfo>
    <mods:language>
      <mods:languageTerm  authority="iso639-2b"
        type="code">eng</mods:languageTerm>
    </mods:language>
    <mods:physicalDescription>
      <mods:form  authority="marcform">electronic</mods:form>
      <mods:digitalOrigin>born digital</mods:digitalOrigin>
    </mods:physicalDescription>
<!-- ... -->
```

```
  </mods:mods>
</xenoData>
```

*Example* This example shows GeoJSON embedded in <xenoData>. Note that JSON does not permit newlines inside string values. These must be escaped as `\n`. To avoid the accidental insertion of newlines by software, the use of *xml:space* is recommended. Blocks of JSON should be wrapped in CDATA sections, as they may contain characters illegal in XML.

```
<xenoData xml:space="preserve">
<![CDATA[
{
    "features": [
        {
            "geometry": {
                "type": "Point",
                "coordinates": [
                    68.388483,
                    33.498616
                ]
            },
            "type": "Feature",
            "id": "darmc-location-19727",
            "properties": {
                "snippet": "Unknown; 330 BC - AD 300",
                "link": "https://pleiades.stoa.org/places/59694/darmc-location-19727",
                "description": "5M scale point location",
                "location_precision": "precise",
                "title": "DARMC location 19727"
            }
        }
    ],
    "id": "59694",
    "subject": [
        "dare:ancient=1",
        "dare:feature=settlement",
        "dare:major=0"
    ],
    "title": "Arachosiorum Oppidum/Alexandria",
    "provenance": "Barrington Atlas: BAtlas 6 B3 Arachosiorum Oppidum/Alexandria",
    "placeTypeURIs": [
        "https://pleiades.stoa.org/vocabularies/place-types/settlement"
    ],
    "details": "<p>The Barrington Atlas Directory notes: Kalat-e-Ghilzai? AFG</p>",
    "@context": {
        "snippet": "dcterms:abstract",
        "rights": "dcterms:rights",
        "description": "dcterms:description",
        "title": "dcterms:title",
        "dcterms": "http://purl.org/dc/terms/",
        "subject": "dcterms:subject",
        "uri": "@id",
        "created": "dcterms:created"
    },
    "review_state": "published",
    "type": "FeatureCollection",
    "description": "An ancient place, cited: BAtlas 6 B3 Arachosiorum Oppidum/Alexandria",
    "reprPoint": [
```

```
            68.388483,
            33.498616
        ],
        "placeTypes": [
            "settlement"
        ],
        "bbox": [
            68.388483,
            33.498616,
            68.388483,
            33.498616
        ],
        "rights": "Copyright © The Contributors. Sharing and remixing permitted under CC-BY.",
        "created": "2010-09-23T04:31:55Z",
        "uri": "https://pleiades.stoa.org/places/59694",
        "creators": [
            {
                "username": null,
                "name": "M.U. Erdosy"
            }
        ],
        "@type": "Place"
    }
    ]]>
    </xenoData>
```

Note: the example above has been trimmed for legibility. The original may be found linked from Arachosiorum Oppidum/Alexandria. The contributors, listed per the license terms, are R. Talbert, Jeffrey Becker, W. Röllig, Tom Elliott, H. Kopp, DARMC, Sean Gillies, B. Siewert-Mayer, Francis Deblauwe, and Eric Kansa.

*Schematron* <sch:pattern is-a="declarable"> <sch:param name="tde" value="tei:xenoData"/> </sch:pattern>

*Content model*

```
<content>
 <alternate>
  <textNode/>
  <anyElement minOccurs="1"
   maxOccurs="unbounded"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element xenoData
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.declarable.attributes,
   att.typed.attributes,
```

```
   ( text | anyElement_xenoData_1+ )
}
```

---

**<xr>** (cross-reference phrase) contains a phrase, sentence, or icon referring the reader to some other location in this or another text. [*10.3.5.3. Cross-References to Other Entries*]

*Module* dictionaries — Dictionaries

*Attributes*
- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*

  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*
  - att.global.responsibility
    - *@cert*
    - *@resp*
  - att.global.source
    - *@source*
- att.lexicographic
  - *@expand*
  - *@split*
  - *@value*
  - *@location*

- *@mergedIn*
- *@opt*

- att.datcat
  - *@datcat*
  - *@valueDatcat*
  - *@targetDatcat*
- att.lexicographic.normalized
  - *@norm*
  - *@orig*
- att.typed
  - ~~type~~
  - @subtype

`@type` indicates the type of cross reference, using any convenient typology.

*Derived from* att.typed

*Status* Recommended

*Datatype* teidata.enumerated

*Sample values include:* **syn** (synonym) cross reference for synonym information

**etym** (etymological) etymological information

**cf** (compare or consult) related or similar term

**illus** (illustration) illustration of an object

*Member of* model.entryPart model.entryPart.top

*Contained by*

*core:* cit

*dictionaries:* dictScrap entry entryFree etym hom re sense

*namesdates:* nym

*May contain*

*analysis:* c cl interp interpGrp m pc phr s span spanGrp w

*certainty:* certainty precision respons

*core:* abbr add address bibl biblStruct binaryObject cb choice cit corr date del desc distinct ellipsis email emph expan foreign gap gb gloss graphic hi index label lb list listBibl measure measureGrp media mentioned milestone name note noteGrp num orig pb ptr q quote ref reg rs ruby said sic soCalled stage term time title unclear unit

*dictionaries:* lang lbl oRef pRef usg

*drama:* camera caption castList move sound tech view

*figures:* figure formula notatedMusic table

*gaiji:* g

*header:* biblFull idno

*iso-fs:* fLib fs fvLib

*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline

*msdescription:* catchwords depth dim dimensions height heraldry locus locusGrp material msDesc objectType origDate origPlace secFol signatures stamp watermark width

*namesdates:* addName affiliation bloc climate country district eventName forename genName geo geogFeat geogName listEvent listNym listObject listOrg listPerson listPlace listRelation location nameLink objectName offset orgName persName persPronouns placeName population region roleName settlement state surname terrain trait

*spoken:* incident kinesic pause shift vocal writing

*tagdocs:* att classSpec code constraintSpec dataSpec eg egXML elementSpec gi ident listRef macroSpec moduleSpec outputRendition specDesc specGrp specGrpRef specList tag val

*textcrit:* app listApp listWit witDetail

*textstructure:* floatingText

*transcr:* addSpan am damage damageSpan delSpan ex fw handShift listTranspose metamark mod redo restore retrace secl space subst substJoin supplied surplus undo

*verse:* caesura rhyme

character data

*Note* May contain character data and phrase-level elements; usually contains a <ref> or a <ptr> element.

This element encloses both the actual indication of the location referred to, which may be tagged using the <ref> or <ptr> elements, and any accompanying material which gives more information about why the reader is being referred there.

*Example*

```
<entry>
  <form>
    <orth>lavage</orth>
  </form>
  <etym>[Fr. < <mentioned>laver</mentioned>; L. <mentioned>lavare</mentioned>,
to wash;
  <xr>see <ref>lather</ref>
    </xr>]. </etym>
</entry>
```

*Example*

```
<entry>
  <form>
    <orth>lawful</orth>
  </form>
  <xr  type="syn">SYN. see <ref>legal</ref>
  </xr>
</entry>
```

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <elementRef key="usg"/>
  <elementRef key="lbl"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element xr
{
   att.global.attributes,
```

```
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.lexicographic.attributes,
   att.datcat.attributes,
   att.lexicographic.normalized.attributes,
   att.typed.attribute.subtype,
   attribute type { text }?,
   (
      text
    | model.gLike    | model.phrase    | model.inter    | usg    | lbl    | model.global
}
```

---

**<zone>** defines any two-dimensional area within a <surface> element. [*12.1. Digital Facsimiles 12.2.2. Embedded Transcription*]

*Module* transcr — Representation of Primary Sources

*Attributes*

- att.global
  - *@xml:id*
  - *@n*
  - *@xml:lang*
  - *@xml:base*
  - *@xml:space*
  - att.global.analytic
    - *@ana*
  - att.global.change
    - *@change*
  - att.global.facs
    - *@facs*
  - att.global.linking
    - *@corresp*
    - *@synch*
    - *@sameAs*
    - *@copyOf*
    - *@next*
    - *@prev*
    - *@exclude*
    - *@select*
  - att.global.rendition
    - *@rend*
    - *@style*
    - *@rendition*

  - att.global.responsibility
    * *@cert*
    * *@resp*
  - att.global.source
    * *@source*
- att.coordinated
  - *@start*
  - *@ulx*
  - *@uly*
  - *@lrx*
  - *@lry*
  - *@points*
- att.typed
  - *@type*
  - *@subtype*
- att.written
  - *@hand*

`@rotate` indicates the amount by which this zone has been rotated clockwise, with respect to the normal orientation of the parent <surface> element as implied by the dimensions given in the <msDesc> element or by the coordinates of the <surface> itself. The orientation is expressed in arc degrees.
*Status* Optional
*Datatype* teidata.numeric
*Default* 0

*Member of* model.linePart model.standOffPart
*Contained by*
*linking:* standOff
*transcr:* line surface zone
*May contain*
*analysis:* c interp interpGrp pc span spanGrp w
*certainty:* certainty precision respons
*core:* add binaryObject cb choice del ellipsis gap gb graphic hi index lb media milestone note noteGrp pb q unclear
*figures:* figure formula notatedMusic
*gaiji:* g
*iso-fs:* fLib fs fvLib
*linking:* alt altGrp anchor join joinGrp link linkGrp seg timeline
*spoken:* incident kinesic pause shift vocal writing
*textcrit:* app witDetail
*transcr:* addSpan damage damageSpan delSpan fw handShift line listTranspose metamark mod path redo restore retrace space substJoin surface undo zone
character data
*Note* The position of every zone for a given surface is always defined by reference to the coordinate system defined for that surface.

A graphic element contained by a zone represents the whole of the zone.

A zone may be of any shape. The attribute *points* may be used to define a polygonal zone, using the coordinate system defined by its parent surface.

A zone is always a closed polygon. Repeating the initial coordinate at the end of the sequence is optional. To encode an unclosed path, use the <path> element.

*Example*

```
<surface ulx="14.54" uly="16.14" lrx="0"
  lry="0">
  <graphic url="stone.jpg"/>
  <zone points="4.6,6.3 5.25,5.85 6.2,6.6 8.19222,7.4125 9.89222,6.5875
10.9422,6.1375 11.4422,6.7125 8.21722,8.3125 6.2,7.65"/>
</surface>
```

This example defines a non-rectangular zone: see the illustration in section *12.2.2.1. Advanced Uses of surface and zone*.

*Example*

```
<facsimile>
  <surface ulx="50" uly="20" lrx="400"
    lry="280">
    <zone ulx="0" uly="0" lrx="500" lry="321">
      <graphic url="graphic.png"/>
    </zone>
  </surface>
</facsimile>
```

This example defines a zone which has been defined as larger than its parent surface in order to match the dimensions of the graphic it contains.

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.graphicLike"/>
  <classRef key="model.global"/>
  <elementRef key="surface"/>
  <classRef key="model.linePart"/>
 </alternate>
</content>
```

*Schema Declaration*

```
element zone
{
   att.global.attributes,
   att.global.analytic.attributes,
   att.global.change.attributes,
   att.global.facs.attributes,
   att.global.linking.attributes,
   att.global.rendition.attributes,
   att.global.responsibility.attributes,
   att.global.source.attributes,
   att.coordinated.attributes,
   att.typed.attributes,
   att.written.attributes,
```

```
   attribute rotate { text }?,
   (
      text
    | model.gLike    | model.graphicLike    | model.global    | surface    | model.linePar
}
```

# D Attributes

## D.1 About the Attributes Appendix

This appendix gives you a list of attributes and links to the reference pages for the elements on which they appear. There are 275 distinctly-named attributes in revision ebf0834e1 of TEI P5 Version 4.10.0 of the TEI Guidelines.

| | |
|---|---|
| absolute | when |
| active | interaction relation |
| adj | node |
| adjFrom | node |
| adjTo | node |
| age | person personGrp persona |
| agent | att.damaged gap unclear |
| aloud | said |
| ana | att.global.analytic |
| anchored | att.anchoring |
| arity | tree |
| assertedValue | certainty |
| atLeast | att.ranging |
| atMost | att.ranging |
| attachment | surface |
| atts | specDesc |
| autoPrefix | content |
| baseForm | m |
| baseTypes | fsDecl |
| behaviour | model |
| break | att.breaking |
| calendar | att.calendarSystem |
| cause | att.textCritical att.transcriptional |
| cert | att.global.responsibility certainty |
| change | att.global.change |
| children | iNode root |
| class | att.msClass attRef |
| code | occupation socecStatus |
| cols | att.tableDecoration table |
| columns | layout |
| commodity | att.measurement |
| confidence | att.ranging |
| contemporary | binding seal |
| copyOf | att.global.linking |
| corresp | att.global.linking |
| cRef | att.cReferencing |
| cssClass | model |
| datcat | att.datcat |
| datingMethod | att.datable.custom |
| datingPoint | att.datable.custom |
| datum | geoDecl |
| decls | att.declaring |
| default | att.declarable |
| defaultExceptions | schemaSpec |
| defective | att.msExcerpt |
| degree | att.damaged certainty node purpose |
| delim | citeStructure refState |
| dim | space |

| | |
|---|---|
| direct | said |
| discrete | sound |
| docLang | schemaSpec |
| domains | att.pointing.group |
| dur | att.duration.w3c |
| dur-iso | att.duration.iso |
| ed | att.edition |
| edRef | att.edition |
| encoding | binaryObject |
| end | att.timed |
| enjamb | att.enjamb |
| eol | hyphenation |
| evaluate | att.pointing |
| evidence | att.editLike persPronouns |
| except | anyElement classRef moduleRef |
| exclude | att.global.linking |
| expand | att.lexicographic classRef |
| extent | att.dimensions att.partials |
| facs | att.global.facs |
| feats | fs |
| feature | shift |
| filter | equiv |
| flipping | surface |
| follow | iNode leaf |
| force | pc |
| form | objectDesc |
| formula | att.formula |
| from | att.citing att.datable.w3c app arc locus rt span |
| from-custom | att.datable.custom |
| from-iso | att.datable.iso |
| fromUnit | conversion |
| full | att.personal |
| function | att.segLike material metamark |
| fVal | f |
| gender | actor person personGrp persona role |
| generate | classSpec |
| generatedBy | att.cmc |
| gi | tagUsage |
| given | certainty |
| gradual | writing |
| group | att.damaged |
| hand | att.written |
| hands | handDesc |
| height | att.media |
| ident | att.identified application language prefixDef remarks transcriptionDesc valItem |
| include | classRef moduleRef |
| inDegree | node |
| indentLevel | att.indentation |
| indexName | index |
| inst | att.interpLike |
| instant | att.editLike |
| interval | timeline when |
| iterated | kinesic vocal |
| join | att.linguistic |

| | |
|---|---|
| key | att.canonical classRef dataRef elementRef macroRef memberOf moduleRef schemaRef specDesc |
| label | rhyme |
| lang | code |
| lemma | att.linguistic |
| lemmaRef | att.linguistic |
| length | refState |
| level | langKnown sense title |
| loc | app |
| location | att.lexicographic variantEncoding |
| locus | certainty respons |
| lrx | att.coordinated |
| lry | att.coordinated |
| mainLang | textLang |
| marks | punctuation quotation |
| match | att.scoping citeStructure |
| matchPattern | att.patternReplacement |
| material | supportDesc |
| max | att.ranging memberOf numeric |
| maxOccurs | att.repeatable datatype |
| medium | att.handFeatures |
| mergedIn | att.lexicographic |
| met | att.metrical |
| method | correction normalization variantEncoding |
| mimeType | att.internetMedia media |
| min | att.ranging memberOf |
| minOccurs | att.repeatable datatype |
| modality | post |
| mode | att.combinable alt altGrp channel classes memberOf |
| module | att.identified |
| motivation | annotation |
| msd | att.linguistic |
| mutual | relation |
| n | att.global |
| name | att.gaijiProp attRef dataFacet dataRef equiv f fDecl namespace param relation unicodeProp unihanProp vLabel |
| new | handShift shift |
| next | att.global.linking |
| norm | att.lexicographic.normalized |
| notAfter | att.datable.w3c |
| notAfter-custom | att.datable.custom |
| notAfter-iso | att.datable.iso |
| notation | att.notated |
| notBefore | att.datable.w3c |
| notBefore-custom | att.datable.custom |
| notBefore-iso | att.datable.iso |
| ns | att.namespaceable |
| nymRef | att.naming |
| occurs | tagUsage |
| opt | att.lexicographic |
| optional | fDecl |
| ord | iNode root tree |
| order | graph tree |
| ordered | listChange |

| | |
|---|---|
| org | att.divLike attList vColl vMerge |
| orig | att.lexicographic.normalized |
| origin | timeline |
| otherLangs | textLang |
| outDegree | iNode node root |
| output | model modelGrp modelSequence |
| parent | iNode leaf |
| part | att.fragmentable |
| partial | tagsDecl |
| parts | nym |
| passive | interaction relation |
| pattern | metDecl |
| perf | move tech |
| period | att.datable |
| place | att.placement |
| placement | punctuation |
| points | att.coordinated path |
| pos | att.linguistic |
| pre | pc |
| precision | att.dimensions precision |
| predeclare | att.identified |
| predicate | att.predicate modelSequence |
| prefix | elementSpec moduleRef schemaSpec |
| preserveOrder | sequence |
| prev | att.global.linking |
| property | citeData |
| quantity | att.dimensions att.measurement |
| queryBinding | constraintDecl |
| real | att.metrical |
| reason | gap secl supplied surplus unclear |
| ref | att.canonical dataRef g |
| rend | att.global.rendition |
| rendition | att.global.rendition |
| replacementPattern | att.patternReplacement |
| replyTo | post |
| require | att.textCritical anyElement |
| resp | att.global.responsibility space |
| restriction | dataRef |
| result | join joinGrp |
| rhyme | att.metrical |
| role | att.naming att.tableDecoration org person personGrp persona |
| rotate | zone |
| rows | att.tableDecoration table |
| ruledLines | layout |
| sameAs | att.global.linking |
| sample | att.divLike |
| scale | att.media |
| scheme | att.gaijiProp att.styleDef att catRef classCode constraintDecl constraintSpec gi keywords locus locusGrp occupation socecStatus tag |
| schemeVersion | att.styleDef |
| scope | att.dimensions att.scope join outputRendition rendition |
| scribe | att.handFeatures |
| scribeRef | att.handFeatures |
| script | att.handFeatures |

| | |
|---|---|
| scriptRef | att.handFeatures |
| select | att.global.linking |
| selector | rendition |
| seq | att.transcriptional |
| sex | actor person personGrp persona |
| since | when |
| size | graph personGrp |
| social | distinct |
| sort | att.personal |
| sortKey | att.sortable |
| source | att.global.source |
| space | distinct |
| spanTo | att.spanning |
| split | att.lexicographic |
| start | att.coordinated att.timed schemaSpec |
| status | att.docStatus att.transcriptional availability correction |
| stdDeviation | precision |
| streams | layout |
| style | att.global.rendition |
| subtype | att.interpLike att.typed |
| synch | att.global.linking |
| tag | langKnown |
| tags | langKnowledge |
| target | att.pointing att.scoping alt annotation change fsdLink material metamark redo relatedItem rt specGrpRef undo |
| targetDatcat | att.datcat |
| targetEnd | att.anchoring |
| targetLang | att.pointing schemaSpec |
| targFunc | att.pointing.group |
| terminal | metSym |
| time | distinct |
| to | att.citing att.datable.w3c app arc locus rt span |
| to-custom | att.datable.custom |
| to-iso | att.datable.iso |
| toUnit | conversion |
| toWhom | att.ascribed.directed |
| trans | u |
| trunc | numeric |
| type | att.entryLike att.interpLike att.textCritical att.typed abbr affiliation age app birth castItem certainty classSpec constitution constraintSpec correspAction death derivation desc dimensions distinct divGen domain education factuality faith form fs fsDecl fsdLink fw gram graph iType idno interaction langKnowledge lbl list listForest measure metDecl move nationality node num oRef occupation orth preparedness purpose q recording residence socecStatus sound stage tag tech title titlePage titlePart usg valList witDetail xr |
| ulx | att.coordinated |
| uly | att.coordinated |
| unit | att.citing att.dimensions att.measurement att.milestoneUnit citeStructure pc timeline when |
| unitRef | att.measurement |
| uri | equiv |
| url | att.resourced moduleRef |
| usage | attDef language |
| use | att.citeStructurePart |
| useSourceRendition | model modelGrp modelSequence |

| | |
|---|---|
| valid | egXML |
| validUntil | att.deprecated |
| value | att.gaijiProp att.lexicographic age binary dataFacet eLeaf eTree gender iNode leaf metSym node num numeric param persPronouns root sex symbol triangle unicodeProp unihanProp |
| valueDatcat | att.datcat |
| varSeq | att.textCritical |
| version | att.gaijiProp TEI application teiCorpus transcriptionDesc |
| versionDate | att.translatable |
| weights | alt |
| when | att.datable.w3c |
| when-custom | att.datable.custom |
| when-iso | att.datable.iso |
| where | att.locatable move |
| who | att.ascribed |
| width | att.media |
| wit | att.rdgPart att.witnessed witDetail |
| withId | tagUsage |
| writtenLines | layout |
| xml:base | att.global |
| xml:id | att.global annotation |
| xml:lang | att.global |
| xml:space | att.global |

# E Datatypes and Other Macros

## E.1 About the Datatypes and Macros Appendix

This appendix gives you a list of datypes and links to the reference pages for them. There are 35 distinctly-named data specifications in revision ebf0834e1 of TEI P5 Version 4.10.0 of the TEI Guidelines.

---

**macro.abContent** (anonymous block content) defines the content of anonymous block elements. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* ab

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.paraPart"/>
  <elementRef key="ab"/>
 </alternate>
</content>
```

*Declaration*

```
macro.abContent = ( text | model.paraPart | ab )*
```

---

**macro.limitedContent** (paragraph content) defines the content of prose elements that are not used for transcription of extant materials. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* desc fDescr figDesc fsDescr meeting rendition tagUsage

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.limitedPhrase"/>
  <classRef key="model.inter"/>
 </alternate>
</content>
```

*Declaration*

```
macro.limitedContent = ( text | model.limitedPhrase | model.inter )*
```

---

**macro.paraContent** (paragraph content) defines the content of paragraphs and similar elements. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* add camera caption case colloc corr damage def del docEdition emph gen gram hi hyph iType imprimatur lang lbl mod mood number orig orth p per pos pron ref reg restore retrace rhyme salute secl seg sic signed sound stress subc supplied surplus syll tech title titlePart tns u unclear usg writing

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.paraPart"/>
 </alternate>
</content>
```

*Declaration* `macro.paraContent = ( text | model.paraPart )*`

---

## macro.phraseSeq (phrase sequence) defines a sequence of character data and phrase-level elements. [*1.4.1. Standard Content Models*]

*Module* tei — The TEI Infrastructure

*Used by* abbr actor addName addrLine affiliation author biblScope birth bloc catchwords citedRange cl colophon country death distinct distributor district docAuthor docDate edition editor education eg email eventName expan explicit extent faith finalRubric floruit foreign forename fw genName gender geoDecl geogFeat geogName gloss headItem headLabel heraldry incipit label material measure mentioned name nameLink nationality num objectName objectType offset orgName origPlace persName persPronouns phr placeName pubPlace publisher rb region residence role roleDesc roleName rs rt rubric s secFol settlement sex soCalled socecStatus speaker stamp street surname term unit watermark wit

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.attributable"/>
  <classRef key="model.phrase"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Declaration*

```
macro.phraseSeq =
   ( text | model.gLike | model.attributable | model.phrase | model.global )*
```

---

## macro.phraseSeq.limited (limited phrase sequence) defines a sequence of character data and those phrase-level elements that are not typically used for transcribing extant documents. [*1.4.1. Standard Content Models*]

*Module* tei — The TEI Infrastructure

*Used by* activity age authority channel classCode collection constitution derivation domain factuality funder institution interaction langKnown language locale metSym preparedness principal purpose repository resp span sponsor valDesc

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.limitedPhrase"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Declaration*

```
macro.phraseSeq.limited = ( text | model.limitedPhrase | model.global )*
```

---

**macro.specialPara** ('special' paragraph content) defines the content model of elements such as notes or list items, which either contain a series of component-level elements or else have the same structure as a paragraph, containing a series of phrase-level and inter-level elements. [*1.3. The TEI Class System*]

*Module* tei — The TEI Infrastructure

*Used by* accMat acquisition additions cell change collation condition custEvent decoNote filiation foliation handNote item layout licence metamark musicNotation note occupation origin provenance q quote said scriptNote signatures source stage summary support surrogates textLang typeNote view

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.divPart"/>
  <classRef key="model.global"/>
 </alternate>
</content>
```

*Declaration*

```
macro.specialPara =
   (
      text
    | model.gLike    | model.phrase    | model.inter    | model.divPart    | model.global
```

---

**macro.specialPara.cmc** ('special' paragraph content with added CMC element(s)) defines the content model of the <quote> element, whose content is just like others (such as notes or list items) which either contain a series of component-level elements or else have the same structure as a paragraph, containing a series of phrase-level and inter-level elements, but also contains element(s) for transcription of CMC.

*Module* cmc — Computer-mediated Communication

*Used by*

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
  <classRef key="model.phrase"/>
  <classRef key="model.inter"/>
  <classRef key="model.divPart"/>
  <classRef key="model.global"/>
  <classRef key="model.cmc"/>
 </alternate>
</content>
```

*Declaration*

```
macro.specialPara.cmc =
   (
      text
    | model.gLike    | model.phrase    | model.inter    | model.divPart    | model.global
```

---

## macro.xtext (extended text) defines a sequence of character data and gaiji elements.

*Module* tei — The TEI Infrastructure

*Used by* c depth dim ex height mapping memberOf string width

*Content model*

```
<content>
 <alternate minOccurs="0"
  maxOccurs="unbounded">
  <textNode/>
  <classRef key="model.gLike"/>
 </alternate>
</content>
```

*Declaration*

```
macro.xtext = ( text | model.gLike )*
```

---

## teidata.authority defines attribute values which derive from an authority list, which may be an enumerated list defined in the document's schema, a list or taxonomy elsewhere in the document, or an online taxonomy, gazetteer, or other authority.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- move/@where

*Content model*

```
<content>
 <alternate>
```

```
  <dataRef key="teidata.enumerated"/>
  <dataRef key="teidata.pointer"/>
 </alternate>
</content>
```

*Declaration*

```
teidata.authority = teidata.enumerated | teidata.pointer
```

*Note* Attribute values with this datatype should either come from a value list in the attribute specification (`teidata.enumerated`) or be a valid URI (`teidata.pointer`).

---

**teidata.certainty** defines the range of attribute values expressing a degree of certainty.

*Module* tei — The TEI Infrastructure

*Used by* teidata.probCertElement:

- certainty/@cert
- precision/@precision
- purpose/@degree

*Content model*

```
<content>
 <valList type="closed">
  <valItem ident="high"/>
  <valItem ident="medium"/>
  <valItem ident="low"/>
  <valItem ident="unknown"/>
 </valList>
</content>
```

*Declaration* `teidata.certainty = "high" | "medium" | "low" | "unknown"`

*Note* Certainty may be expressed by one of the predefined symbolic values high, medium, or low. The value unknown should be used in cases where the encoder does not wish to assert an opinion about the matter.

---

**teidata.count** defines the range of attribute values used for a non-negative integer value used as a count.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- age/@value
- datatype/@minOccurs
- graph/@order
- graph/@size
- handDesc/@hands
- iNode/@outDegree
- layout/@columns
- layout/@streams

- layout/@ruledLines
- layout/@writtenLines
- node/@inDegree
- node/@outDegree
- node/@degree
- refState/@length
- root/@outDegree
- sense/@level
- table/@rows
- table/@cols
- tagUsage/@occurs
- tagUsage/@withId
- tree/@arity
- tree/@order

*Content model*

```
<content>
 <dataRef name="nonNegativeInteger"/>
</content>
```

*Declaration* `teidata.count = xsd:nonNegativeInteger`

*Note* Any positive integer value or zero is permitted

---

**teidata.duration.iso** defines the range of attribute values available for representation of a duration in time using ISO 8601 standard formats.

*Module* tei — The TEI Infrastructure

*Used by*

*Content model*

```
<content>
 <dataRef name="token"
  restriction="[0-9.,DHMPRSTWYZ/:+\-]+"/>
</content>
```

*Declaration*

```
teidata.duration.iso = token { pattern = "[0-9.,DHMPRSTWYZ/:+\-]+" }
```

*Example*

```
<time dur-iso="PT0,75H">three-quarters of an hour</time>
```

*Example*

```
<date dur-iso="P1,5D">a day and a half</date>
```

*Example*

```
<date dur-iso="P14D">a fortnight</date>
```

*Example*

```
<time dur-iso="PT0.02S">20 ms</time>
```

*Note* A duration is expressed as a sequence of number-letter pairs, preceded by the letter P; the letter gives the unit and may be Y (year), M (month), D (day), H (hour), M (minute), or S (second), in that order. The numbers are all unsigned integers, except for the last, which may have a decimal component (using either `.` or `,` as the decimal point; the latter is preferred). If any number is *0*, then that number-letter pair may be omitted. If any of the H (hour), M (minute), or S (second) number-letter pairs are present, then the separator `T` must precede the first 'time' number-letter pair.

For complete details, see ISO 8601 *Data elements and interchange formats — Information interchange — Representation of dates and times*.

---

**teidata.duration.w3c** defines the range of attribute values available for representation of a duration in time using W3C datatypes.

*Module* tei — The TEI Infrastructure

*Used by*

*Content model* `<content> <dataRef name="duration"/></content>`

*Declaration* `teidata.duration.w3c = xsd:duration`

*Example*

```
<time dur="PT45M">forty-five minutes</time>
```

*Example*

```
<date dur="P1DT12H">a day and a half</date>
```

*Example*

```
<date dur="P7D">a week</date>
```

*Example*

```
<time dur="PT0.02S">20 ms</time>
```

*Note* A duration is expressed as a sequence of number-letter pairs, preceded by the letter P; the letter gives the unit and may be Y (year), M (month), D (day), H (hour), M (minute), or S (second), in that order. The numbers are all unsigned integers, except for the `S` number, which may have a decimal component (using `.` as the decimal point). If any number is *0*, then that number-letter pair may be omitted. If any of the H (hour), M (minute), or S (second) number-letter pairs are present, then the separator `T` must precede the first 'time' number-letter pair.

For complete details, see the W3C specification.

---

**teidata.enumerated** defines the range of attribute values expressed as a single XML name taken from a list of documented possibilities.

*Module* tei — The TEI Infrastructure

*Used by* teidata.authority teidata.gender teidata.sexElement:

- abbr/@type
- affiliation/@type
- age/@type
- alt/@mode
- altGrp/@mode
- annotation/@motivation
- app/@type
- att/@scheme
- attDef/@usage
- attList/@org
- availability/@status
- birth/@type
- castItem/@type
- certainty/@type
- certainty/@locus
- channel/@mode
- citeStructure/@unit
- classSpec/@type
- classSpec/@generate
- classes/@mode
- constitution/@type
- constraintDecl/@scheme
- constraintDecl/@queryBinding
- constraintSpec/@type
- constraintSpec/@scheme
- correction/@status
- correction/@method
- correspAction/@type
- death/@type
- derivation/@type
- desc/@type
- dimensions/@type
- distinct/@type
- divGen/@type
- domain/@type
- education/@type
- egXML/@valid
- factuality/@type
- faith/@type
- form/@type

- fs/@type
- fsDecl/@type
- fsdLink/@type
- fw/@type
- gap/@reason
- gap/@agent
- geoDecl/@datum
- gi/@scheme
- gram/@type
- graph/@type
- hyphenation/@eol
- iType/@type
- idno/@type
- interaction/@type
- interaction/@active
- interaction/@passive
- join/@scope
- langKnowledge/@type
- lbl/@type
- list/@type
- listForest/@type
- material/@function
- measure/@type
- memberOf/@mode
- metDecl/@type
- model/@behaviour
- model/@output
- modelGrp/@output
- modelSequence/@output
- move/@type
- nationality/@type
- node/@type
- normalization/@method
- num/@type
- oRef/@type
- objectDesc/@form
- occupation/@type
- org/@role
- orth/@type
- outputRendition/@scope

- param/@name
- pc/@force
- pc/@unit
- persPronouns/@evidence
- persPronouns/@value
- person/@role
- person/@age
- personGrp/@role
- personGrp/@age
- persona/@role
- persona/@age
- post/@modality
- preparedness/@type
- punctuation/@marks
- punctuation/@placement
- purpose/@type
- q/@type
- quotation/@marks
- recording/@type
- relation/@name
- rendition/@scope
- residence/@type
- respons/@locus
- secl/@reason
- shift/@feature
- shift/@new
- socecStatus/@type
- sound/@type
- space/@dim
- stage/@type
- supplied/@reason
- supportDesc/@material
- surface/@attachment
- tag/@type
- tag/@scheme
- tech/@type
- timeline/@unit
- title/@type
- title/@level
- titlePage/@type

- titlePart/@type
- tree/@ord
- u/@trans
- unclear/@reason
- unclear/@agent
- usg/@type
- vColl/@org
- vMerge/@org
- valList/@type
- variantEncoding/@method
- variantEncoding/@location
- when/@unit
- witDetail/@type
- xr/@type

*Content model* `<content> <dataRef key="teidata.word"/></content>`

*Declaration* `teidata.enumerated = teidata.word`

*Note* Attributes using this datatype must contain a single 'word' which contains only letters, digits, punctuation characters, or symbols: thus it cannot include whitespace.

Typically, the list of documented possibilities will be provided (or exemplified) by a value list in the associated attribute specification, expressed with a <valList> element.

---

**teidata.gender** defines the range of attribute values used to represent the gender of a person, persona, or character.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- actor/@gender
- gender/@value
- person/@gender
- personGrp/@gender
- persona/@gender
- role/@gender

*Content model*

```
<content>
 <dataRef key="teidata.enumerated"/>
</content>
```

*Declaration* `teidata.gender = teidata.enumerated`

*Note* Values for attributes using this datatype may be defined locally by a project, or they may refer to an external standard.

Values for this datatype should not be used to encode morphological gender (cf. <gen>, *msd* as defined in `att.linguistic`, and *10.3.1. Information on Written and Spoken Forms*).

**teidata.interval** defines attribute values used to express an interval value.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- timeline/@interval
- when/@interval

*Content model*

```
<content>
 <alternate>
  <dataRef name="float"/>
  <valList>
   <valItem ident="regular"/>
   <valItem ident="irregular"/>
   <valItem ident="unknown"/>
  </valList>
 </alternate>
</content>
```

*Declaration*

```
teidata.interval = xsd:float | ( "regular" | "irregular" | "unknown" )
```

*Note* Any value greater than zero or any one of the values regular, irregular, unknown.

---

**teidata.language** defines the range of attribute values used to identify a particular combination of human language and writing system. [*vi.1 Language Identification*]

*Module* tei — The TEI Infrastructure

*Used by* Element:

- langKnowledge/@tags
- langKnown/@tag
- language/@ident
- schemaSpec/@targetLang
- schemaSpec/@docLang
- textLang/@mainLang
- textLang/@otherLangs

*Content model*

```
<content>
 <alternate>
  <dataRef name="language"/>
  <valList>
   <valItem ident=""/>
  </valList>
 </alternate>
</content>
```

*Declaration*

```
teidata.language = xsd:language | ( "" )
```

*Note* The values for this attribute are language 'tags' as defined in BCP 47. Currently BCP 47 comprises RFC 5646 and RFC 4647; over time, other IETF documents may succeed these as the best current practice.

A 'language tag', per BCP 47, is assembled from a sequence of components or *subtags* separated by the hyphen character (-, U+002D). The tag is made of the following subtags, in the following order. Every subtag except the first is optional. If present, each occurs only once, except the fourth and fifth components (variant and extension), which are repeatable.

**language** The IANA-registered code for the language. This is almost always the same as the ISO 639 2-letter language code if there is one. The list of available registered language subtags can be found at `https://www.iana.org/assignments/language-subtag-registry`. It is recommended that this code be written in lower case.

**script** The ISO 15924 code for the script. These codes consist of 4 letters, and it is recommended they be written with an initial capital, the other three letters in lower case. The canonical list of codes is maintained by the Unicode Consortium, and is available at `https://unicode.org/iso15924/iso15924-codes.html`. The IETF recommends this code be omitted unless it is necessary to make a distinction you need.

**region** Either an ISO 3166 country code or a UN M.49 region code that is registered with IANA (not all such codes are registered, e.g. UN codes for economic groupings or codes for countries for which there is already an ISO 3166 2-letter code are not registered). The former consist of 2 letters, and it is recommended they be written in upper case; the list of codes can be searched or browsed at `https://www.iso.org/obp/ui/#search/code/`. The latter consist of 3 digits; the list of codes can be found at `http://unstats.un.org/unsd/methods/m49/m49.htm`.

**variant** An IANA-registered variation. These codes 'are used to indicate additional, well-recognized variations that define a language or its dialects that are not covered by other available subtags'.

**extension** An extension has the format of a single letter followed by a hyphen followed by additional subtags. There are currently only two extensions in use. Extension `T` indicates that the content was transformed. For example en-t-it could be used for content in English that was translated from Italian. Extension T is described in the informational RFC 6497. Extension `U` can be used to embed a variety of locale attributes. It is described in the informational RFC 6067.

**private use** An extension that uses the initial subtag of the single letter *x* (i.e., starts with `x-`) has no meaning except as negotiated among the parties involved. These should be used with great care, since they interfere with the interoperability that use of RFC 4646 is intended to promote. In order for a document that makes use of these subtags to be TEI-conformant, a corresponding <language> element must be present in the TEI header.

There are two exceptions to the above format. First, there are language tags in the IANA registry that do not match the above syntax, but are present because they have been 'grandfathered' from previous specifications.

Second, an entire language tag can consist of only a private use subtag. These tags start with `x-`, and do not need to follow any further rules established by the IETF and endorsed by these Guidelines. Like all language tags that make use of private use subtags, the language in question must be documented in a corresponding <language> element in the TEI header.

Examples include

**sn** Shona

**zh-TW** Taiwanese

**zh-Hant-HK** Chinese written in traditional script as used in Hong Kong

**en-SL** English as spoken in Sierra Leone

**pl** Polish

**es-MX** Spanish as spoken in Mexico

**es-419** Spanish as spoken in Latin America

The W3C Internationalization Activity has published a useful introduction to BCP 47, Language tags in HTML and XML.

---

## teidata.name defines the range of attribute values expressed as an XML Name.

*Module* tei — The TEI Infrastructure

*Used by* att giElement:

- application/@ident
- attRef/@class
- attRef/@name
- equiv/@name
- f/@name
- fDecl/@name
- fsDecl/@baseTypes
- index/@indexName
- join/@result
- joinGrp/@result
- memberOf/@key
- model/@cssClass
- remarks/@ident
- schemaSpec/@start
- specDesc/@key
- specDesc/@atts
- tagUsage/@gi
- transcriptionDesc/@ident

*Content model* `<content> <dataRef name="Name"/></content>`

*Declaration* `teidata.name = xsd:Name`

*Note* Attributes using this datatype must contain a single word which follows the rules defining a legal XML name (see `https://www.w3.org/TR/REC-xml/#dt-name`): for example they cannot include whitespace or begin with digits.

---

## teidata.namespace defines the range of attribute values used to indicate XML namespaces as defined by the W3C Namespaces in XML Technical Recommendation.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- anyElement/@require
- namespace/@name

*Content model*

```
<content>
 <dataRef restriction="\S+" name="anyURI"/>
</content>
```

*Declaration* `teidata.namespace = xsd:anyURI { pattern = "\S+" }`

*Note* The range of syntactically valid values is defined by RFC 3986 *Uniform Resource Identifier (URI): Generic Syntax*

---

**teidata.namespaceOrName** defines attribute values which contain either an absolute namespace URI or a qualified XML name.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- anyElement/@except
- schemaSpec/@defaultExceptions

*Content model*

```
<content>
 <alternate>
  <dataRef name="anyURI">
   <dataFacet name="pattern"
    value="[^/\s]+:\S*"/>
  </dataRef>
  <dataRef name="Name">
   <dataFacet name="pattern" value=".+:.+"/>
  </dataRef>
 </alternate>
</content>
```

*Declaration* `teidata.namespaceOrName = xsd:anyURI | xsd:Name`

---

**teidata.nullOrName** defines attribute values which contain either the null string or an XML name.

*Module* tei — The TEI Infrastructure

*Used by*

*Content model*

```
<content>
 <alternate>
  <valList>
   <valItem ident=""/>
  </valList>
  <dataRef name="NCName"/>
 </alternate>
</content>
```

*Declaration* `teidata.nullOrName = ( "" ) | xsd:NCName`

*Note* The rules defining an XML name form a part of the XML Specification.

---

**teidata.numeric** defines the range of attribute values used for numeric values.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- memberOf/@max
- memberOf/@min
- num/@value
- numeric/@value
- numeric/@max
- precision/@stdDeviation
- zone/@rotate

*Content model*

```
<content>
 <alternate>
  <dataRef name="double"/>
  <dataRef name="token"
   restriction="(\-?[\d]+/\-?[\d]+)"/>
  <dataRef name="decimal"/>
 </alternate>
</content>
```

*Declaration*

```
teidata.numeric =
   xsd:double | token { pattern = "(\-?[\d]+/\-?[\d]+)" } | xsd:decimal
```

*Note* Any numeric value, represented as a decimal number, in floating point format, or as a ratio.

To represent a floating point number, expressed in scientific notation, 'E notation', a variant of 'exponential notation', may be used. In this format, the value is expressed as two numbers separated by the letter E. The first number, the significand (sometimes called the mantissa) is given in decimal format, while the second is an integer. The value is obtained by multiplying the mantissa by 10 the number of times indicated by the integer. Thus the value represented in decimal notation as 1000.0 might be represented in scientific notation as 10E3.

A value expressed as a ratio is represented by two integer values separated by a solidus (/) character. Thus, the value represented in decimal notation as 0.5 might be represented as a ratio by the string 1/2.

---

## teidata.outputMeasurement

**teidata.outputMeasurement** defines a range of values for use in specifying the size of an object that is intended for display.

*Module* tei — The TEI Infrastructure

*Used by*

*Content model*

```
<content>
 <dataRef name="token"
  restriction="[\-+]?\d+(\.\d+)?(%|cm|mm|in|pt|pc|px|em|ex|ch|rem|vw|vh|vmin|vmax)"/>
</content>
```

*Declaration*

```
teidata.outputMeasurement =
   token
```

```
   {
      pattern = "[\-+]?\d+(\.\d+)?(%|cm|mm|in|pt|pc|px|em|ex|ch|rem|vw|vh|vmin|vmax)"
   }
```

*Example*

```
<figure>
  <head>The TEI Logo</head>
  <figDesc>Stylized yellow angle brackets with the letters
<mentioned>TEI</mentioned> in
     between and <mentioned>text encoding initiative</mentioned> underneath,
all on a white
     background.</figDesc>
  <graphic  height="600px"  width="600px"
    url="http://www.tei-c.org/logos/TEI-600.jpg"/>
</figure>
```

*Note* These values map directly onto the values used by XSL-FO and CSS. For definitions of the units see those specifications; at the time of this writing the most complete list is in the CSS3 working draft.

---

## teidata.pattern defines attribute values which are expressed as a regular expression.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- dataRef/@restriction
- metDecl/@pattern

*Content model* `<content> <dataRef name="token"/></content>`

*Declaration* `teidata.pattern = token`

*Note* A regular expression, often called a *pattern*, is an expression that describes a set of strings. They are usually used to give a concise description of a set, without having to list all elements. For example, the set containing the three strings *Handel*, *Händel*, and *Haendel* can be described by the pattern `H(ä|ae?)ndel` (or alternatively, it is said that the pattern `H(ä|ae?)ndel` *matches* each of the three strings)Wikipedia

This TEI datatype is mapped to the XSD token datatype, and may therefore contain any string of characters. However, it is recommended that the value used conform to the particular flavour of regular expression syntax supported by XSD Schema.

---

## teidata.point defines the data type used to express a point in cartesian space.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- path/@points

*Content model*

```
<content>
 <dataRef name="token"
  restriction="(-?[0-9]+(\.[0-9]+)?,-?[0-9]+(\.[0-9]+)?)"/>
</content>
```

*Declaration*

```
teidata.point = token { pattern = "(-?[0-9]+(\.[0-9]+)?,-?[0-9]+(\.[0-9]+)?)" }
```

*Example*

```
<facsimile>
  <surface ulx="0" uly="0" lrx="400" lry="280">
    <zone points="220,100 300,210 170,250 123,234">
      <graphic url="handwriting.png"/>
    </zone>
  </surface>
</facsimile>
```

*Note* A point is defined by two numeric values, which should be expressed as decimal numbers. Neither number can end in a decimal point. E.g., both 0.0,84.2 and 0,84 are allowed, but 0.,84. is not.

---

## teidata.pointer

**teidata.pointer** defines the range of attribute values used to provide a single URI, absolute or relative, pointing to some other resource, either within the current document or elsewhere.

*Module* tei — The TEI Infrastructure

*Used by* teidata.authorityElement:

- alt/@target
- annotation/@target
- app/@from
- app/@to
- arc/@from
- arc/@to
- catRef/@scheme
- certainty/@given
- change/@target
- citeData/@property
- classCode/@scheme
- conversion/@fromUnit
- conversion/@toUnit
- dataRef/@ref
- eLeaf/@value
- eTree/@value
- equiv/@uri
- equiv/@filter
- f/@fVal
- fs/@feats
- fsdLink/@target
- g/@ref
- handShift/@new
- iNode/@value

- iNode/@children
- iNode/@parent
- iNode/@follow
- keywords/@scheme
- leaf/@value
- leaf/@parent
- leaf/@follow
- locus/@scheme
- locusGrp/@scheme
- material/@target
- metamark/@target
- moduleRef/@url
- move/@perf
- node/@value
- node/@adjTo
- node/@adjFrom
- node/@adj
- nym/@parts
- occupation/@scheme
- occupation/@code
- post/@replyTo
- redo/@target
- relatedItem/@target
- relation/@active
- relation/@mutual
- relation/@passive
- root/@value
- root/@children
- rt/@target
- rt/@from
- rt/@to
- socecStatus/@scheme
- socecStatus/@code
- space/@resp
- span/@from
- span/@to
- specGrpRef/@target
- tech/@perf
- timeline/@origin
- triangle/@value

- undo/@target
- when/@since
- witDetail/@wit

*Content model*

```
<content>
 <dataRef restriction="\S+" name="anyURI"/>
</content>
```

*Declaration* `teidata.pointer = xsd:anyURI { pattern = "\S+" }`

*Note* The range of syntactically valid values is defined by RFC 3986 *Uniform Resource Identifier (URI): Generic Syntax*. Note that the values themselves are encoded using RFC 3987 *Internationalized Resource Identifiers* (IRIs) mapping to URIs. For example,
`https://secure.wikimedia.org/wikipedia/en/wiki/%` is encoded as
`https://secure.wikimedia.org/wikipedia/en/wiki/%25` while
`http://مصر.الاتصالات-وزارة.موقع/` is encoded as
`http://xn--4gbrim.xn----rmckbbajlc6dj7bxne2c.xn--wgbh1c/`

---

## teidata.prefix defines a range of values that may function as a URI scheme name.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- prefixDef/@ident

*Content model*

```
<content>
 <dataRef name="token"
  restriction="[a-z][a-z0-9\+\.\-]*"/>
</content>
```

*Declaration*

```
teidata.prefix = token { pattern = "[a-z][a-z0-9\+\.\-]*" }
```

*Note* This datatype is used to constrain a string of characters to one that can be used as a URI scheme name according to RFC 3986, section 3.1. Thus only the 26 lowercase letters a–z, the 10 digits 0–9, the plus sign, the period, and the hyphen are permitted, and the value must start with a letter.

---

## teidata.probability defines the range of attribute values expressing a probability.

*Module* tei — The TEI Infrastructure

*Used by* teidata.probCertElement:

- alt/@weights
- certainty/@degree

*Content model*

```
<content>
 <dataRef name="double">
  <dataFacet name="minInclusive" value="0"/>
```

```
  <dataFacet name="maxInclusive" value="1"/>
 </dataRef>
</content>
```

*Declaration* `teidata.probability = xsd:double`

*Note* Probability is expressed as a real number between 0 and 1; 0 representing *certainly false* and 1 representing *certainly true*.

---

**teidata.probCert** defines a range of attribute values which can be expressed either as a numeric probability or as a coded certainty value.

*Module* tei — The TEI Infrastructure

*Used by*

*Content model*

```
<content>
 <alternate>
  <dataRef key="teidata.probability"/>
  <dataRef key="teidata.certainty"/>
 </alternate>
</content>
```

*Declaration*

```
teidata.probCert = teidata.probability | teidata.certainty
```

---

**teidata.replacement** defines attribute values which contain a replacement template.

*Module* tei — The TEI Infrastructure

*Used by*

*Content model* `<content> <textNode/></content>`

*Declaration* `teidata.replacement = text`

---

**teidata.sex** defines the range of attribute values used to identify the sex of an organism.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- actor/@sex
- person/@sex
- personGrp/@sex
- persona/@sex
- sex/@value

*Content model*

```
<content>
 <dataRef key="teidata.enumerated"/>
```

```
</content>
```

*Declaration* `teidata.sex = teidata.enumerated`

*Note* Values for attributes using this datatype may be defined locally by a project, or they may refer to an external standard.

---

**teidata.temporal.iso** defines the range of attribute values expressing a temporal expression such as a date, a time, or a combination of them, that conform to the international standard *Data elements and interchange formats – Information interchange – Representation of dates and times*.

*Module* tei — The TEI Infrastructure

*Used by*

*Content model*

```
<content>
 <alternate>
  <dataRef name="date"/>
  <dataRef name="gYear"/>
  <dataRef name="gMonth"/>
  <dataRef name="gDay"/>
  <dataRef name="gYearMonth"/>
  <dataRef name="gMonthDay"/>
  <dataRef name="time"/>
  <dataRef name="dateTime"/>
  <dataRef name="token"
   restriction="[0-9.,DHMPRSTWYZ/:+\-]+"/>
 </alternate>
</content>
```

*Declaration*

```
teidata.temporal.iso =
   xsd:date
 | xsd:gYear
 | xsd:gMonth
 | xsd:gDay
 | xsd:gYearMonth
 | xsd:gMonthDay
 | xsd:time
 | xsd:dateTime
 | token { pattern = "[0-9.,DHMPRSTWYZ/:+\-]+" }
```

*Note* If it is likely that the value used is to be compared with another, then a time zone indicator should always be included, and only the dateTime representation should be used.

For all representations for which ISO 8601:2004 describes both a *basic* and an *extended* format, these Guidelines recommend use of the extended format.

---

**teidata.temporal.w3c** defines the range of attribute values expressing a temporal expression such as a date, a time, or a combination of them, that conform to the W3C XML Schema Part 2: Datatypes Second Edition specification.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- when/@absolute

*Content model*

```
<content>
 <alternate>
  <dataRef name="date"/>
  <dataRef name="gYear"/>
  <dataRef name="gMonth"/>
  <dataRef name="gDay"/>
  <dataRef name="gYearMonth"/>
  <dataRef name="gMonthDay"/>
  <dataRef name="time"/>
  <dataRef name="dateTime"/>
 </alternate>
</content>
```

*Declaration*

```
teidata.temporal.w3c =
   xsd:date
 | xsd:gYear
 | xsd:gMonth
 | xsd:gDay
 | xsd:gYearMonth
 | xsd:gMonthDay
 | xsd:time
 | xsd:dateTime
```

*Note* If it is likely that the value used is to be compared with another, then a time zone indicator should always be included, and only the dateTime representation should be used.

---

**teidata.temporal.working** defines the range of values, conforming to the W3C XML Schema Part 2: Datatypes Second Edition specification, expressing a date or a date and a time within the working life of the document.

*Module* tei — The TEI Infrastructure

*Used by*

*Content model*

```
<content>
 <alternate>
  <dataRef name="date"
   restriction="(19[789][0-9]|[2-9][0-9]{3}).*"/>
  <dataRef name="dateTime"
   restriction="(19[789][0-9]|[2-9][0-9]{3}).*"/>
 </alternate>
</content>
```

*Declaration*

```
teidata.temporal.working =
   xsd:date { pattern = "(19[789][0-9]|[2-9][0-9]{3}).*" }
 | xsd:dateTime { pattern = "(19[789][0-9]|[2-9][0-9]{3}).*" }
```

*Note* If it is likely that the value used is to be compared with another, then a time zone indicator should always be included, and only the dateTime representation should be used.

The earliest time expressable with this datatype is 01 January 1970 (the Unix Epoch), which could be written as either 1970-01-01 or 1970-01-01T00:00:00Z.

---

**teidata.text** defines the range of attribute values used to express some kind of identifying string as a single sequence of Unicode characters possibly including whitespace.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- distinct/@time
- distinct/@space
- distinct/@social
- refState/@delim
- rendition/@selector
- unicodeProp/@value
- valItem/@ident

*Content model* `<content> <dataRef name="string"/></content>`

*Declaration* `teidata.text = string`

*Note* Attributes using this datatype must contain a single 'token' in which whitespace and other punctuation characters are permitted.

---

**teidata.truthValue** defines the range of attribute values used to express a truth value.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- binary/@value
- content/@autoPrefix
- fDecl/@optional
- listChange/@ordered
- metSym/@terminal
- model/@useSourceRendition
- modelGrp/@useSourceRendition
- modelSequence/@useSourceRendition
- numeric/@trunc
- pc/@pre
- sequence/@preserveOrder
- surface/@flipping
- tagsDecl/@partial

*Content model* `<content> <dataRef name="boolean"/></content>`

*Declaration* `teidata.truthValue = xsd:boolean`

*Note* The possible values of this datatype are 1 or true, or 0 or false.

This datatype applies only for cases where uncertainty is inappropriate; if the attribute concerned may have a value other than true or false, e.g. unknown, or inapplicable, it should have the extended version of this datatype: `teidata.xTruthValue`.

---

**teidata.unboundedCount** defines the range of values used for a counting number or the string unbounded for infinity.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- datatype/@maxOccurs

*Content model*

```
<content>
 <alternate>
  <dataRef name="nonNegativeInteger"/>
  <valList type="closed">
   <valItem ident="unbounded"/>
  </valList>
 </alternate>
</content>
```

*Declaration*

```
teidata.unboundedCount = xsd:nonNegativeInteger | ( "unbounded" )
```

---

**teidata.version** defines the range of attribute values which may be used to specify a TEI or Unicode version number.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- TEI/@version
- teiCorpus/@version

*Content model*

```
<content>
 <dataRef name="token"
  restriction="[\d]+(\.[\d]+){0,2}"/>
</content>
```

*Declaration*

```
teidata.version = token { pattern = "[\d]+(\.[\d]+){0,2}" }
```

*Note* The value of this attribute follows the pattern specified by the Unicode consortium for its version number (`https://unicode.org/versions/`). A version number contains digits and fullstop characters only. The first number supplied identifies the major version number. A second and third number, for minor and sub-minor version numbers, may also be supplied.

---

## teidata.versionNumber defines the range of attribute values used for version numbers.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- application/@version
- transcriptionDesc/@version

*Content model*

```
<content>
 <dataRef name="token"
  restriction="[\d]+[a-z]*[\d]*(\.[\d]+[a-z]*[\d]*){0,3}"/>
</content>
```

*Declaration*

```
teidata.versionNumber =
   token { pattern = "[\d]+[a-z]*[\d]*(\.[\d]+[a-z]*[\d]*){0,3}" }
```

---

## teidata.word defines the range of attribute values expressed as a single word or token.

*Module* tei — The TEI Infrastructure

*Used by* teidata.enumeratedElement:

- app/@loc
- binaryObject/@encoding
- certainty/@assertedValue
- code/@lang
- dataFacet/@name
- langKnown/@level
- locus/@from
- locus/@to
- m/@baseForm
- media/@mimeType
- metSym/@value
- metamark/@function
- personGrp/@size
- rhyme/@label
- surplus/@reason
- symbol/@value
- unihanProp/@value
- vLabel/@name

*Content model*

```
<content>
 <dataRef name="token"
  restriction="[^\p{C}\p{Z}]+"/>
```

```
</content>
```

*Declaration* `teidata.word = token { pattern = "[^\p{C}\p{Z}]+" }`

*Note* Attributes using this datatype must contain a single 'word' which contains only letters, digits, punctuation characters, or symbols: thus it cannot include whitespace.

---

## teidata.xmlName defines attribute values which contain an XML name.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- classRef/@key
- classRef/@include
- classRef/@except
- dataRef/@key
- dataRef/@name
- elementRef/@key
- elementSpec/@prefix
- macroRef/@key
- moduleRef/@prefix
- moduleRef/@include
- moduleRef/@except
- moduleRef/@key
- schemaRef/@key
- schemaSpec/@prefix
- unicodeProp/@name
- unihanProp/@name

*Content model* `<content> <dataRef name="NCName"/></content>`

*Declaration* `teidata.xmlName = xsd:NCName`

*Note* The rules defining an XML name form a part of the XML Specification.

---

## teidata.xpath defines attribute values which contain an XPath expression.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- citeStructure/@match
- modelSequence/@predicate
- param/@value

*Content model* `<content> <textNode/></content>`

*Declaration* `teidata.xpath = text`

*Note* Any XPath expression using the syntax defined in **?**.

When writing programs that evaluate XPath expressions, programmers should be mindful of the possibility of malicious code injection attacks. For further information about XPath injection attacks, see the article at OWASP.

---

## teidata.xTruthValue (extended truth value) defines the range of attribute values used to express a truth value which may be unknown.

*Module* tei — The TEI Infrastructure

*Used by* Element:

- binding/@contemporary
- iNode/@ord
- kinesic/@iterated
- root/@ord
- said/@aloud
- said/@direct
- seal/@contemporary
- sound/@discrete
- vocal/@iterated
- writing/@gradual

*Content model*

```
<content>
 <alternate>
  <dataRef name="boolean"/>
  <valList>
   <valItem ident="unknown"/>
   <valItem ident="inapplicable"/>
  </valList>
 </alternate>
</content>
```

*Declaration*

```
teidata.xTruthValue = xsd:boolean | ( "unknown" | "inapplicable" )
```

*Note* In cases where where uncertainty is inappropriate, use the datatype `teidata.TruthValue`.

# F Bibliography

## F.1 Works Cited in Examples in these Guidelines


[1] 阿城，《棋王》。

[2] Académie française, *Rectifications de l'orthographe - J.O. du 06-12-1990*, En ligne, consulté le 05-03-2010.

[3] Adams, Douglas. *The Hitchhiker's Guide to the Galaxy*, New York: Pocket Books, 1979, chapter 31.

[4] Agence bibliographique de l'enseignement supérieur, *ABES*:Site internetpar l'ABES,consulté le 05-03-2010.

[5] 阿拉伯短劍，國史館：嚴家淦總統文物。

[6] Alighieri, Dante. 'Doglia mi reca ne lo core ardire', *Rime*, XLIX.

[7] Allinson, E.P. and B. Penrose. *Philadelphia 1681-1887* (1887), p. 138.

[8] *American National Standard for Bibliographic References, ANSI Z39.29-1977*, New York: American National Standards Institute (1977).

[9] Andersson, Theodore M.. *A Preface to the Nibelungenlied*, Stanford University Press (1987).

[10] Andrews, Mr.. 'Song', *Chambers's Edinburgh Journal Series 1* 9 :463 (12 December 1840), 376. https://dvpp.uvic.ca/poems/chambers_series/1840/pom_5242_song.html

[11] Anouilh, Jean, *Antigone*, 1842.

[12] *[As in Visions of]* Single leaf of Notes for a poem about night "visions," possibly related to the untitled 1855 poem that Whitman eventually titled "The Sleepers." Fragments of an unidentified newspaper clipping about the Puget Sound area have been pasted to the leaf. The Trent Collection of Walt Whitman Manuscripts, Duke University Rare Book, Manuscript, and Special Collections Library. http://www.whitmanarchive.org/resources/sleepers/duk.00258.001.jpg

[13] Atkins et al. *Collins Robert French-English English-French Dictionary*. London: Collins (1978).

[14] Atkinson, J. Maxwell and John Heritage. *Structures of social action: Studies in conversation analysis*, Cambridge and Paris: Cambridge University Press, Editions de la Maison des Sciences de l'Homme (1984), ix-xvi.

[15] Austen, Jane. *Pride and Prejudice*. (1813), chapter 1.

[16] 白先勇，〈金大班的最後一夜〉，《台北人》。

[17] 白先勇，《孽子》。

[18] 白居易，《憶江南》。

[19] *Amheida I: Ostraka from Trimithis Volume 1: Texts from the 2004–2007 Seasons*, Bagnall, R. S. and G. R. Ruffini, with contributions by R. Cribiore and G. Vittmann (2012).

[20] Baker, James K.. *Night in Tarras*. In *Hilltop: A Literary Paper*, vol 1 no 2. Wellington: Victoria University College Literary Society. (1949).

[21] Balzac, Honoré de, *Les Chouans*, 1845.

[22] Balzac, Honoré de, Le Père Goriot, 1843.

[23] Balzac, Honoré de, *Petites misères de la vie conjugale*, 1850.

[24] Barbauld, Lucy Aikin. *The Works of Anna Laetitia Barbauld* (1826).

[25] Barker, Jane. *The Lining to the Patch-Work Screen* (1726).

[26] *Base de datos paleográfica da lírica galego-portuguesa (PalMed)*. Versión 1.2. Santiago de Compostela: Centro Ramón Piñeiro para a Investigación en Humanidades. http://www.cirp.gal/palmed. f. B126r, column a, l. 21-32.

[27] Bataille, Michel , *L'Arbre de Noël*, 1967.

[28] 'Les Chats', in : Baudelaire, Charles, *Les Fleurs du mal*, 1861.

[29] 'La Vie antérieure', in : Baudelaire, Charles, *Les Fleurs du mal*, 1861.

[30] 電影《霸王別姬》，1993年。

[31] Beck, Béatrice, *Léon Morin, prêtre*, 1952.

[32] Beckett, Samuel. *Waiting for Godot*, London: Faber and Faber (1956).

[33] Beckett, Samuel. *Murphy* (1963), chap 2.

[34] Becque, Henry, *La Parisienne*. Edition électronique par l'ATILF et le CNRTL, d'après l'édition de Fasquelle (Paris, 1922).

[35] Bowers, Jack *Mixtepec-Mixtec Project Personography* `https://raw.githubusercontent.com/iljackb/Mixtepec_Mixtec/master/MIX-People.xml#element(JS/6)`

[36] Beerbohm, Max. Autograph manuscript of *The Golden Drugget*, Pierpont Morgan MA 3391. in ? 123.

[37] Behn, Aphra. *The Rover*, (1697).

[38] Beeton, Isabella. *The book of Household Management*, London: S.O. Beeton (1861).

[39] Belloc, Denis , *Képas*, 1989.

[40] Belloc, Denis , *Néons*, 1987.

[41] Bentham, Jeremy. *The Book of Fallacies*. (1824).

[42] *Beowulf and The fight at Finnsburg*; edited, with introduction, bibliography, notes, glossary, and appendices, by Fr. Klaeber. Boston, New York [etc.] D.C. Heath & Co. (1922).

[43] *Bibliographie dans le cadre de la semaine italienne du 11 au 18 mars 2006*, , document électronique.

[44] Bibliothèque nationale de France, *Projet de description des reliures remarquables de la Réserve des livres rares selon le modèle de la TEI manuscrits*.

[45] Billetdoux, Marie, *Un peu de désir sinon je meurs*, 2006.

[46] Blake, William. 'London', in *Songs of Experience* (1791).

[47] Blake, William. 'The Sick Rose', in *Songs of Experience* (1794).

[48] Blake, William. 'The Tyger', in *Songs of Experience* (1794).

[49] Bloomfield, Leonard. 'Literate and Illiterate Speech', *American Speech*, 2 , (1927), pp. 432-441.

[50] Borges, Jorge Luis, tr. R. Simms. 'The Analytical Language of John Wilkins'. In Emir Rodriguez Monegal and Alistair Reid, eds. *Borges: A reader*, Dutton Adult (1981), p.141.

[51] Borges, Jorge Luis. 'Avatars of the Tortoise' In James E. Irby tr. *Labyrinths: Selected Stories and Other Writings*, New York: New Directions, (1962), pp.202-203.

[52] Bouillier, Grégoire, *Rapport sur moi*, 2002.

[53] Bresson, Robert, 'Mouchette : script', *l'Avant-scène cinéma*, n° 80, avril 1968.

[54] Extract from *British National Corpus* (`http://www.natcorp.ox.ac.uk`) Text KB7, sentence 13730.

[55] Brontë, Charlotte. *Jane Eyre: An Autobiography*, Third edition; reprintLondon Service & Paton, 1897; Project Gutenberg, 1 December 2020. chapter XII. `https://www.gutenberg.org/files/1260/1260-h/1260-h.htm`

[56] Browning, Robert. *Letter to George Moulton-Barrett*, Pierpont Morgan MA 310, (? 23).

[57] Buirette de Belloy, Pierre Laurent, *Gabrielle de Vergy*, 1777.

[58] Bunyan, John. *The Pilgrim's Progress from this world to that which is to come…*, London (1678).

[59] Burgess, Anthony. *A Clockwork Orange*. (1962), opening.

[60] 'Burnard SGML reading list', *COMP.TEXT.SGML*, September 30 to October 11, 1990 in *Usenet Archives* (UsenetArchives.com, 2024).

[61] Burnard, Lou. 'Principles of Database Design' in S. Rahtz. ed. *Information Technology in the Humanities: tools, techniques and applications*, Ellis Horwood Ltd, Ellis Horwood Series in Computers and Their Applications, (1987), p. 54.

[62] Burnard, Lou, Sperberg-McQueen, C. M., *La TEI simplifiée, une introduction au codage des textes électroniques en vue de leur échange — version de travail*, 1996.

[63] Burton, Robert. *Anatomy of Melancholy* (1621), 16th ed. reprinted 1846, p. 743.

[64] Butler, Samuel. *The Way of All Flesh* (1903), chapter 37.

[65] Byron, George Gordon. *Don Juan* (1819), I.xxii.

[66] Byron, George Gordon. 'Vision of Judgment' In E.H. Coleridge ed. *The Poetical Works of Lord Byron*, viii, 1922.

[67] *C 60/16 Fine Roll 6 HENRY III (28 October 1221-27 October 1222)*, membrane 5, entry 154.

[68] CBETA

[69] *Cædmon's Hymn* in Bede's Historia Ecclesiastica (MS Kk. 5. 16, Cambridge, University Library).

[70] *Cabaret*. A musical play, with book by Joe Masteroff, lyrics by Fred Ebb, and music by John Kaner. Based on the play by John van Druten and stories by Christopher Isherwood. 1966

[71] Edward Barkley, describing how Essex drove the Irish from the plains into the woods to freeze or famish in winter; quoted by Canny, Nicholas P. 'The Ideology of English Colonization: From Ireland to America'. In Stanley N. Katz and John M. Murrin eds. *Colonial America: Essays in Politics and Social Development*, 3d ed New York: Knopf, (1983), p.53.

[72] Carroll, Lewis. *Through the Looking Glass, and what Alice found there*. (1871).

[73] *Catéchisme de l'Eglise catholique*, 1968.

[74] Cavendish, Margaret. 'Nature's Pictures'. London, 1656. Women Writers Online. Women Writers Project, Northeastern University. 29 Mar. 2015.

[75] Centre national de la recherche scientifique (France). UMR 7118 ATILF, *Le Trésor de la Langue Française Informatisé (TLFI)*, 2004.

[76] Example recoded from Chafe, W. 'Adequacy, user-friendliness, and practicality in transcribing' In Leech, G., G. Myers, J. Thomas eds. *Spoken English on Computer: Transcription, Markup and Applications*. Harlow: Longman, 1995.

[77] Chandler, Lloyd. 'Conversation with Death' (also known as 'Oh, Death'). In *Journal of Folklore Research*, 41.2/3 , (2004), pp. 125-126.

[78] Chapman, Graham, Cleese, John, Gilliam, Terry, Idle, Eric, Jones, Terry. *The complete Monty Pythons Flying Circus*.

[79] Chapman, Graham, Cleese, John, Gilliam, Terry, Idle, Eric, Jones, Terry. *Monty Python's Life of Brian* (1979).

[80] Chaucer, Geoffrey. *Canterbury Tales*, f52r, in Holkham MS.

[81] Chaucer, Geoffrey. 'The Tale of Sir Topas', *The Canterbury Tales*, In F. N. Robinson ed. *The Works of Geoffrey Chaucer*, 2nd edition Boston: Houghton Mifflin Co., 1957.

[82] 陳政彥，〈戰後臺灣現代詩論戰史研究〉，2007。

[83] Cholières, Nicolas de, *La Forest Nuptiale* (1600).

[84] Chomsky, Noam and Morris Halle. *The Sound Pattern of English*. New York: Harper & Row (1968), p. 415.

[85] Claude Simon *Le jardin des plantes*, 1997, p 284

[86] Cleaver, Eldridge. *Soul on Ice*. New York (1968).

[87] 'Cloud of Unknowing' In Hodgson, Phyllis ed. *The Cloud of Unknowing and The Book of Privy Counselling*, London: Oxford University Press, Early English Text Society, 218, (1944).

[88] Clover, Carol J. *The Medieval Saga*, Ithaca: Cornell University Press (1982).

[89] Cocteau, Jean. *La Machine Infernale*.

[90] Coleridge, Samuel Taylor. *The Rime of the Ancient Mariner*. In Wordsworth, William and Samuel Taylor Coleridge. *Lyrical Ballads* (1798).

[91] Coleridge, Samuel Taylor. 'Frost at Midnight' In E.H. Coleridge ed. *Poetical Works*, Oxford: Oxford University Press, (1967), p.240.

[92] Colette, *Colette à l'école*, 1900.

[93] *Collins English Dictionary*, 12th edition Glasgow: Collins (2014).

[94] *Collins Pocket Dictionary of the English language*. London: Collins.

[95] Collins, Wilkie. *The Moonstone*, Penguin, 6th narrative.

[96] Comenius, John Amos. *Orbis Pictus: a facsimile of the first English edition of 1659* (ed. John E. Sadler) Oxford University Press (1968).

[97] Cope, Thomas Pym. *Philadelphia merchant: the diary of Thomas P. Cope, 1800-1851*, Eliza Cope Harrison ed.

[98] Corneille, Pierre, *La Place Royale ou L'Amoureux extravagant*, 1637.

[99] Corneille, Pierre, *Théodore, vierge et martyre : tragédie chrétienne*, *1646*.

[100] Cowley, Hannah. *The Runaway* (1813).

[101] Crashaw, Richard, ed. J.R. Tutin. *The Poems of Richard Crashaw*. Muses Library Edition: (1900).

[102] Creeley, Robert *A counterpoint* in *For Love: Poems 1950-1960* (1962).

[103] 崔西．雪佛蘭。《戴珍珠耳環的少女》。台北：皇冠，2003。

[104] Dallas, George Mifflin. *Unpublished letter* cited in Russell F. Weigley, Nicholas B. Wainwright, Edwin Wolf eds. *Philadelphia: A 300 Year History*, New York and London: W. W. Norton & Company, 1982, p. 349.

[105] Danhauser, Adolphe-Leopold, *Théorie de la musique*, 1872.

[106] Darwin, Charles. *Narrative of the Surveying Voyages of His Majesty's ships Adventure and Beagle... volume 3 : Journal and Remarks (The Voyage of the Beagle)*, chap 3. 1839.

[107] Daudet, Alphonse, *Les contes du lundi*, 1873.

[108] Davenant, William. *The vvitts. A comedie, presented at the private house in Blacke Fryers, by his Majesties servants. The author VVilliam D'avenant, servant to Her Majestie.* London, 1636. [STC S109311]

[109] *De Nutrimento et Nutribili, Tractatus 1*, fol 217r col b of Merton College Oxford MS O.2.1; in **?** pl. 16.

[110] *Dean of Sarum Churchwardens' presentments*, 1731, Hurst; Wiltshire Record Office; transcribed by Donald A. Spaeth.

[111] Defoe, Daniel. *The Fortunes and Misfortunes of the Famous Moll Flanders* (1722).

[112] Defoe, Daniel. *Robinson Crusoe* (1719).

[113] Defoe, Daniel. *Journal of the Plague Year*. London (1722).

[114] Dekker, Thomas and Thomas Middleton. *The Honest Whore, Part One* (1604).

[115] Deloney, Thomas. *Thomas of Reading or the Sixe Worthie Yeomen of the West* (1612).

[116] Dennery, Annie. 'Du mélos à la note : Les notations musicales au Moyen Age', *Médiévales*, n° 3 1983.

[117] Dickens, Charles. *A Christmas Carol in Prose, Being a Ghost Story of Christmas*, Chapman and Hall, (1843), p. 5, p. 12.

[118] Dickens, Charles. *Little Dorrit*, (1857).

[119] Dickinson, Emily. '1755' In Arthur Eastman et al. eds. *The Norton Anthology of Poetry*, New York: W.W. Norton, 1970, p.859.

[120] Diderot, Denis, *Lettres à Sophie Volland*, 26 sept. 1762.

[121] 〈第三屆第一次大會第一次會議記錄〉，臺灣省諮議會。

[122] Disraeli, Benjamin. *Coningsby* (1844), preface.

[123] Doyle, Arthur Conan. 'The Red-headed league'. In *The Adventures of Sherlock Holmes* (1892).

[124] Doyle, Arthur Conan. *The Original Illustrated Sherlock Holmes*, Castle Books, 1989.

[125] *Drawing of a leaden plaque bearing an inquiry by Hermon from the oracular precinct at Dodona*. Catalogue no 725; LSAG 230.13. Image from http://poinikastas.csad.ox.ac.uk/

[126] Dubois, Jacques, Nyssen, Hubert (dir. de collection), *Babel*, Actes sud,1989-....

[127] Dubois, Jean, *Dictionnaire de linguistique*, 1974.

[128] Dudo of St Quentin. *De moribus et actis primorum Normannie ducum*, fol 4v of British Library MS Harley 3742; in ? pl 6(i).

[129] Dufournaud, Nicole, *Numérisation des Titres de famille*, édition électronique, 2002.

[130] Duhamel, Georges, *Chronique des Pasquier*, 1933-1945.

[131] Dutt, Toru. 'Lakshman' In *Ancient Ballads and Legends of Hindustan*, Kegan Paul, Trench & Co., (1882)

[132] Dylan, Bob. 'All Along the Watchtower' In *John Wesley Harding* (1967).

[133] Dylan, Bob. 'Positively 4th Street' In *Highway 61 Revisited* (1965).

[134] *Eddic poems*, in Reykjavík, Landsbókasafn Íslands, Lbs 1562 4to.

[135] 'Editorial', *.EXE magazine*, 6.11 (1992), p. 2.

[136] Eliot, George. *Middlemarch* (1871), I.1.

[137] Eliot, George. *Middlemarch* (1871), I.XXI.

[138] Eliot, George. *Daniel Deronda* (1876), III.1.

[139] Eliot, Thomas Stearns. *The waste land: a facsimile and transcript of the original drafts including the annotations of Ezra Pound*, Eliot, Valerie ed. Faber and Faber Ltd. (1971), p. 37.

[140] *Scalable Vector Graphics (SVG) 1.1 (Second Edition)*. Editors Erik Dahlström, Jon Ferraiolo, 藤沢 淳, Anthony Grasso, Dean Jackson, Chris Lilley, Cameron McCormack Doug Schepers, Jonathan Watt and Patrick Dengler. World Wide Web Consortium (W3C) (22 June 2010). Available from https://www.w3.org/TR/SVG11/.

[141] Ernaux, Annie, *Se perdre*, 1988.

[142] Ernaux, Annie, *L'usage de la photo*, 2005.

[143] *Études littéraires*, consulté le 07-05-2010.

[144] Fallet , René, *Le Triporteur*, 1951.

[145] Field, Syd *Screenplay: the Foundations of Screenwriting* (1998).

[146] Fielding, Henry. 'The History of the Adventures of Joseph Andrews and his Friend, Mr. Abraham Abrams' (1742).

[147] Fielding, Henry. *Tragedy of Tragedies*, (1737).

[148] Fish, Stanley. *Is there a text in this class? The authority of interpretive communities*. Harvard University Press (1980).

[149] Fisher, M. F. K. 'I Was Really Very Hungry' In *As They Were*, Knopf (1982), p. 43.

[150] FitzGerald, Edward. *The Rubáiyát of Omar Khayyám, translated into English Verse* (1859), stanza 25.

[151] Flaubert, Gustave, *Salammbô*, 1862.

[152] Flaubert, Gustave, *La Tentation de saint Antoine*, 1874.

[153] Foley James D., Andries van Dam, Steven K. Feiner and John F. Hughes. *Computer Graphics: Principles and Practice*, 2nd edition Reading: Addison-Wesley, p.259.

[154] Forster, E.M., *Howards End*. Edward Arnold. (1910).

[155] *France inter, Là-bas si j'y suis*, consulté le 07-05-2010.

[156] *France inter, Le téléphone sonne*, consulté le 07-05-2010.

[157] *Franglais et franricain*, consulté le 07-05-2010.

[158] Fry, Christopher, *The firstborn* (1948).

[159] 福春嫁女

[160] Fussell, Paul. *The Norton Book of Travel*. W. W. Norton (1987).

[161] Galilei, Galileo, *Sidereus Nuncius*, Venetiis: Apud Thomam Baglionum, 1610, quoted by Tufte, Edward R., *Envisioning Information*, Cheshire: Graphics Press (1990), p. 97.

[162] Galland, Antoine,*Les Mille et une nuits*, 1704.

[163] Galsworthy, John, *Strife*. In *Plays, vol 1.* (1909).

[164] Gaskell, Elizabeth Cleghorn. *The Grey Woman*, MS.

[165] Gavioli, Laura and Gillian Mansfield. *The PIXI corpora: bookshop encounters in English and Italian*, Bologna: Cooperativa Libraria Universitaria Editrice (1990), p.74.

[166] Gay, John. *The Beggar's Opera* (1728).

[167] Gazdar, Gerald and Mellish, Christopher. *Natural language processing in Prolog*. Addison-Wesley (1989), p.5.

[168] Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag: *Generalized Phrase Structure Grammar*, Harvard University Press (1985).

[169] Gershwin, Ira *By Strauss* (from *An American in Paris*, 1953).

[170] Gianfraco Contini, *Poeti del Duecento*, Milano-Napoli:Ricciardi I, (1960). pp 155-64.

[171] Gibbon, Edward, *The History of the Decline and Fall of the Roman Empire*, (1789), chapter 58.

[172] Gide, André, *Le Journal des Faux-Monnayeurs*, 1927.

[173] Gide, André, *Journal : 1889-1939*, 1939.

[174] Gilbert, William Schwenck and Sullivan, Arthur. *HMS Pinafore* (1878), I.

[175] Gilbert, William Schwenck and Sullivan, Arthur. *The Mikado* (1885).

[176] Ginsberg, Allen. 'My alba', *Reality Sandwiches*, San Francisco: City Lights, (1963).

[177] Giono, Jean, *Regain*, 1930.

[178] Giraudoux, Jean, *La guerre de Troie n'aura pas lieu*, 1935.

[179] Godefroy, Frédéric, *Dictionnaire de l'ancienne langue française et de tous ses dialectes du IXe au XVe siècle*, 1881.

[180] Goethe, Johann Wolfgang von, tr. Philip Wayne. *Faust*, Part 1, London: Penguin. (1949).

[181] Goethe, Johann Wolfgang von. *Auf dem See* (1775).

[182] [Goldsmith, Oliver], *The History of Little GOODY TWO-SHOES; Otherwise called, Mrs. MARGERY TWO-SHOES…* (Printed for J. Newbery, 1766).

[183] Grallert, Till, *Tools for the Computational Normalisation and Analysis of Food Prices and Food Riots in the Eastern Mediterranean (Bilād al-Shām) in the 19th and 20th Centuries*. Zenodo, 2021. https://doi.org/10.5281/zenodo.5159020.

[184] Graves, Robert. *Rough draft of letter to Desmond Flower*. 17 Dec 1938. (from *Diary of Robert Graves 1935-39 and ancillary materials*, compiled by Beryl Graves, C.G. Petter, L.R. Roberts, University of Victoria Libraries).

[185] Greene, Robert. *Groatsworth of Wit Bought with a Million of Repentance* (1592).

[186] Gregory of Tours. *Ecclesiastical History of the Franks* (translated by Lewis Thorpe).

[187] Greimas, Algirdas Julien et Courtés, Joseph, *Sémiotique : dictionnaire raisonné de la théorie du langage*, 1979.

[188] Grune, Dick and Ceriel J. H. Jacobs. *Parsing Techniques: A Practical Guide*, New York and London: Ellis Horwood, 1990, p. 24.

[189] Guerard, Françoise. *Le Dictionnaire de Notre Temps*, ed. Paris: Hachette, 1990

[190] Hall, Radclyffe. *The well of loneliness*. (1928).

[191] Halliday, M.A.K. and R. Hassan. *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective*, Oxford: Oxford University Press, 1990, p. 104.

[192] Hanks, Patrick. 'Definitions and Explanations'. In J. M. Sinclair ed. *Looking Up*.Collins, 1987, p. 121.

[193] *Hannam's last farewell to the world:: being a full and true relation of the notorious life and shamfull death of Mr. Richard Hannam...* (1656).

[194] Hansberry, Lorraine. *A raisin in the sun.* (1959).

[195] Harvey, Gabriel. *Four letters and certain sonnets specially touching Robert Greene...* (1592).

[196] Harvey, William. *Exercitatio Anatomica De Motu Cordis Et Sanguinis In Animalibus*, 1949 (Third edition, second printing), p. 74. https://archive.org/details/in.ernet.dli.2015.115852/page/n98/mode/1up

[197] Hobbes, Thomas. *Leviathan* (1681).

[198] 《紅樓夢》第六回：賈寶玉初試云雨情 劉姥姥一進榮國府

[199] 〈紅頭嶼踏查報告〉，國立台灣大學圖書館：田代文庫。

[200] Hoola Van Nooten, Berthe*Fleurs, fruits et feuillages choisis de l'île de Java peints d'après nature*, troisième édition par C. Muquardt, Bruxelles, 1880.

[201] Hornby, A.S. et al. *Oxford Advanced Learner's Dictionary of Current English*. Oxford University Press (1974).

[202] 黃晨淳編著，《希臘羅馬神話故事》，p76。

[203] Hugo, Victor, *Notre-Dame de Paris*, 1845.

[204] 'Souvenir de la nuit du 4', in : Hugo, Victor, *Les Contemplations*, 1853.

[205] Hugo, Victor, *Les Misérables*, 1862.

[206] 'A la fenêtre pendant la nuit', in : Hugo, Victor , *Les Contemplations*, 1856.

[207] Huizinga, Johan, *L'Automne du Moyen-Âge*, 1919.

[208] *ISO 690:1987: Information and documentation – Bibliographic references – Content, form and structure* clause 4.1 , p.2.

[209] *ISO 24611:2012 Language resource management — Morpho-syntactic annotation framework (MAF)*. International Organization for Standardization. 2012.

[210] Ibsen, Henrik, tr. William and Charles Archer. *Peer Gynt* (1875).

[211] Ibsen, Henrik *Den episke Brand* KBK Collin 2869, 4º I, fol. 11v.

[212] Ibsen, Henrik *Digte* NBO Ms.4º 1110a, p. 36.

[213] Ibsen, Henrik *Brand*, The Royal Library, Denmark. KBK Collin 262, 4°, I.1.1, fol. [4]v

[214] Ibsen, Henrik, tr. R. Farquharson Sharp and Eleanor Marx-Aveling. *A Doll's House* In *A Doll's House; and Two Other Plays by Henrik Ibsen*, Everyman's library: the drama 494, London: J. M. Dent & Sons, 1910.

[215] Idle, Eric, Michael Palin, Graham Chapman, John Cleese, Terry Gilliam. *The Complete Monty Python's Flying Circus: All the Words*, Pantheon Books, (1989), 2 , p.230.

[216] Imbs, Paul, Martin, Robert, Straka, Georges, *Mélanges de linguistique française et de philologie et littérature médiévales offerts à Monsieur Paul Imbs*, 1973 .

[217] Institut de Littérature Française Moderne de l'Université de Neuchâtel, *Présentation du projet Cyrus*, En ligne, consulté le 05-03-2010.

[218] Ionesco, Eugène, *La cantatrice chauve*, 1950.

[219] 羅貫中，《三國演義》。

[220] Jakacki, Diane. 'Schema question: further constraining @subtype based on @type'. January 18, 2021. TEI (Text Encoding Initiative) public discussion list, https://listserv.brown.edu/cgi-bin/wa?A2=TEI-L;b092ac32.2101.

[221] Jarry, Alfred, tr. Simon Watson Taylor and Cyril Connolly. *The Ubu plays*, London: Methuen, 1968.

[222] Jean-Jacques Rousseau *Les confessions* vol 6

[223] Jerome, Jerome K. *Three men in a boat* (1889), chapter 6.

[224] Joachim Heinrich Campe *Wörterbuch der Deutschen Sprache. Erster Theil. A - bis - E.*, 1807.

[225] John, Lord Sheffield, *The Auto-biography of Edward Gibbon esq.*, (1846), 222.

[226] Jonson, Ben. *Bartholomew Fair.*

[227] Jonson, Ben. *Volpone*, J. B. Bamborough ed. Macmillan, 1963, p.14.

[228] Jonson, Ben. *The Alchemist*, Douglas Brown ed. London: Benn, 1966, I.1, p.9.

[229] Joyce, James. *Ulysses*: The Bodley Head, 1960. p. 933.

[230] Judith Patt, Michiko Warkentyne and Barry Till. *Haiku: Japanese Art and Poetry.* 2010. Pomegranate Communications, Inc. San Francisco. `http://www.pomegranate.com`.

[231] Karlstadt, Andreas Bodenstein. *Distinctiones Thomistarum.* Wittenberg 1507. In Harald Bollbuck ed. Kritische Gesamtausgabe Karlstadt I, available: `http://diglib.hab.de/edoc/ed000216/start.htm`

[232] Kersey, John. *Dictionarium Anglo-Brittannicum: Or, a General English Dictionary.* London: J. Wilde, 1715. 2nd ed.

[233] Kilian, Wilfrid, *Neige et glaciers ; notes prises au cours de géologie de la Faculté des Sciences de Grenoble par M. Alamelle*, 1891-1895.

[234] King, Martin Luther. *Letter from Birmingham City Jail.* Philadelphia, PA: American Friends Service Committee (1963).

[235] Kipling, Rudyard. 'The mother hive'. In *Actions and Reactions*, London: Macmillan, (1909).

[236] Kipling, Rudyard. *Stalky & Co.*, London: Macmillan (1899).

[237] Kipling, Rudyard. *Kim*, Macmillan, (1901), p. 9.

[238] Knuth, Donald Ervin. 'The Potrzebie System of Weights and Measures', *Mad Magazine*33June 1957.

[239] Koltès, Bernard-Marie,*Quai ouest*, 1985.

[240] Krummrey, Hans, Panciera, Silvio *Criteri di edizione e segni diacritici*, Tituli 2 (1980), 205-215.

[241] Kyd, Thomas. *Spanish Tragedy* (1592).

[242] La Fayette, Marie-Madeleine Pioche de La Vergne, comtesse de, *La Princesse de Clèves*, 1678.

[243] La Fontaine, Jean de. 'L'Astrologue qui se laisse tomber dans un puits' In *Fables Choisies*, Classiques Larousse, Paris: Librairie Larousse, 1, 1940.

[244] La Fontaine, Jean de. 'Le corbeau et le renard' In *Fables Choisies*, Classiques Larousse, Paris: Librairie Larousse, 1, 1940.

[245] Laboratoire ATILF *Analyse et traitement informatique de la langue française (ATILF)*, consulté le 07-05-2010.

[246] Laclos, Pierre Choderlos de. *Les Liaisons dangereuses* (1772), 1963, p. 13.

[247] Ladurie, Emmanuel Le Roy. *Montaillou*, Middlesex: Penguin Books, 1980, p. 3.

[248] Lamartine, Alphonse de, *Méditations poétiques*, 1820.

[249] Lamothe, Alexandre de 'Principaux manuscrits latins et français conservés dans la Bibliothèque impériale et dans celle de l'Ermitage à Saint-Pétersbourg', édition numérique, Bibliothèque de l'école nationale des chartes, 1864.

[250] Langendoen, D. Terence and Paul M. Postal. *Vastness of Natural Languages*, Oxford: Basil Blackwell, 1984, p. 24 , note 12.

[251] Langendoen, D. Terence and Paul Postal. *The vastness of natural languages*, Blackwell, , 1, 1984.

[252] Langland, William. *The Vision of Piers Plowman.* In A.V.C. Schmidt ed. *Langland: Vision of Piers Plowman: "B" Text*, opening.

[253] *Latin liturgical manuscript: Bod MS. Lat liturg. b. 19*, folio 4r. Taken from the digitized image described at `https://www.diamm.ac.uk/sources/522/`.

[254] Lavrentev, Alexei, *Manuel d'encodage XML-TEI étendu des transcriptions de manuscrits dans le projet BFM-Manuscrits*, dernier enregistrement le 26 juin 2008.

[255] Lawrence, David Herbert. Autograph manuscript of *Eloi, Eloi, lama sabachthani*, Pierpont Morgan MA 1892; in ? p.129.

[256] Layamon. *Brut*, fol 65v of Bodleian MS. Rawlinson Poetry 32; in ? 12(ii).

[257] *Le Grand Robert de la langue française*, 2004.

[258] LeTourneau, Mark S. *English Grammar*, 2001, New York: Harcourt p. 89.

[259] Leech, G., G. Myers, J. Thomas eds. *Spoken English on Computer: Transcription, Markup and Applications*. Harlow: Longman, 1995.

[260] Leech, Geoffrey and Mick Short. *Style in Fiction*, London: Longman, 1981, p.272.

[261] Lennon, John and McCartney, Paul. 'She Loves You'. *The Beatles Anthology 1*. Capitol, 1995.

[262] Léry, Jean de, *Histoire faict en la terre de Brésil*, 1580.

[263] Lessing, Doris. *Martha Quest*. 1952, pp. 52-53.

[264] *Letter of Mikołaj Orlik to Mikołaj Serafin, Gródek, January 13 [1438]*. Edited by Skolimowska et al. Available from `https://teipublisher.com/exist/apps/serafin/letters/serafin02.xml`

[265] Lewis, Leopold Davis. *The Bells*, (1871), translated from Erckmann-Chatrian, *Le Juif Polonais*.

[266] Lewis, Wyndham. *Tarr* (1928), Jupiter, 1968, p. 17.

[267] 李白，《黃鶴樓送孟浩然之廣陵》。

[268] 李白，《靜夜思》。

[269] 《梁山伯與祝英台》

[270] 電影《梁山伯與祝英台》，舊版。

[271] 〈歷代漢文大藏經概述〉，李圓淨，原載《南行》第六期（南行學社編印）。

[272] Libelt, Karol. *Wykłady Humboldta na uniwersytecie Berlińskim: notaty prelekcyj tych po uczniu Jego Karolu Libelcie*, Deutsches Textarchiv, 2024-04-02. urn:nbn:de:kobv:b4-30960-7, `https://www.deutschestextarchiv.de/book/download_xml/libelt_hs6623ii_1828`

[273] Lillo, George. *The London Merchant* (1731), epilogue.

[274] 林覺民，《與妻訣別書》。

[275] Lincoln, Abraham. 'Second Inaugural Address', 4 March 1865. In H. S. Commager, ed., *Documents of American History*, 5th ed Crofts American history series. New York: Appleton-Century-Crofts, 1949, p.442.

[276] 劉康，《對話的喧囂：巴赫汀文化理論評述》，台北：麥田，2005，二版。

[277] 劉康，《對話的喧聲：巴赫汀文化理論述評》，台北：麥田，2005，頁20-23。

[278] *Longman Dictionary of Contemporary English*. Harlow, Essex: Longman (1978).

[279] Lorris, Guillaume de et Meun, Jean de*Le Roman de la Rose*, 1230, 1270-1285.

[280] Lowe, David. *Lost Chicago*, Boston: Houghton Mifflin, (1978), p.30.

[281] Lowell, Amy. *Autumn haze*.(1919).

[282] Lucie Lichtig, 'M. Klein de Joseph Losey : script', . *l'Avant-scène cinéma*, n° 175, nov. 1976.

[283] Luther, Martin [tr]. *Die gantze Heilige Schrifft Deudsch, Wittenberg 1545. Letzte zu Luthers Lebzeiten erchienene Ausgabe, hsg. Hans Volz unter Mitarbeit von Heinz Blanke. Textredaktion Friedrich Kur*, München: Rogner & Bernhard, 1972.

[284] 《論語》卷七：憲問第十四。

[285] 羅貫中，《三國演義》。

[286] 魯迅，《狂人日記》。

[287] 魯迅，《狂人日記》。

[288] 魯迅，《狂人日記》。

[289] MacNeice, Louis. 'The Sunlight on the Garden' In E.R. Dodds, *The collected poems of Louis MacNeice*, London: Faber, 1966.

[290] Examples from MacWhinney, Brian, 88, 87, cited by Johansson, S. 'The approach of the Text Encoding Initiative to the encoding of spoken discourse' In Leech, G., G. Myers, J. Thomas eds. *Spoken English on Computer: Transcription, Markup and Applications*. Harlow: Longman, 1995.

[291] Madan, Falconer, et al, *A summary catalogue of western manuscripts in the Bodleian Library at Oxford which have not hitherto been catalogued ...* Oxford, 1895-1953. 5: 515. (Cited in *Driscoll, M.J. P5-MS: A general purpose tagset for manuscript description in Digital Medievalist: 2.1, (2006)*)

[292] Maingeneau, Dominique, *Analyser les textes de communication*, 2007.

[293] Any issue of the *Malawi Daily Times*.

[294] Malle, Louis, 'Zazie dans le métro : script', *l'Avant-scène cinéma*, n° 104, juin 1970.

[295] Manuélian, Hélène, Schang, Emmanuel, *XML, DTD et TEI pour un dictionnaire étymologique des créoles*, édition électronique, 12 Octobre 2007.

[296] 'The Manere of Good Lyuynge': fol. 126v of Bodleian MS Laud Misc 517; in **?**, p.8.

[297] *Marbury v. Madison*, 1 Cranch, 137 (1803), rpt. In H. S. Commager, ed., *Documents of American History*, 5th ed Crofts American history series. New York: Appleton-Century-Crofts, 1949, p.192.

[298] Marvell, Andrew. *An Horatian Ode*, Bod. MS Eng. Poet d.49.

[299] Mauriac, Claude, *La Marquise sortit à cinq heures*, 1961.

[300] Mauss, Marcel, *Sociologie et anthropologie*, 1950.

[301] McCarthy, Niall, 'No Parachute: The Highest Falls People Survived', *Statista*, November 18, 2020. `https://www.statista.com/chart/19708/known-occasions-where-people-survived-falls/`.

[302] McCarty, Willard. 'Introduction' in *Collaborative Research in the Digital Humanities*: A volume in honour of Harold Short on the occasion of his 65th birthday, September 2010, ed. Marilyn Deegan and Willard McCarty. London: Ashgate (2012).

[303] Melville, Herman, *Moby Dick*. (1851).

[304] Mendès-France, Pierre, *Œuvres complètes*, 1984-1985.

[305] 《孟子》〈三十三〉。

[306] Miller, Henry. 'Death of a Salesman' in Atkinson, Brooks, *New Voices in the American Theatre*, New York: Modern Library, 1955, p.113.

[307] Milne, A. A. *The House at Pooh Corner*. London: Methuen & Co., 1928, p. 83.

[308] Milton, John. *Paradise Lost: A poem in X books* (1667), I, 1-10.

[309] Milton, John. *Poems of Mr John Milton, both English and Latin...* (1645).

[310] Molière, *L'École des femmes*,1663.

[311] Molière, *Le Médecin malgré lui*, 1667.

[312] Molière, *Le Médecin volant*, 1673.

[313] Montaigne, Michel de, *Essais*, 1592.

[314] Montesquieu, Charles-Louis de Secondat, baron de la Brède et de, *Lettres persanes*, 1721.

[315] Montherlant, Henry de , *Pitié pour les femmes*, 1936.

[316] Moore, George. Autograph manuscript of *Memoirs of my dead life*, Pierpont Morgan MA 3421; in **?**

[317] Moore, Thomas. Autograph manuscript of the second version of *Lalla Rookh*, Pierpont Morgan MA 310; in **?** 23.

[318] Moreland, Floyd L. and Rita M. Fleischer. *Latin: An Intensive Course,* (1977) p.53.

[319] *Des Minnesangs Frühling*, Moser, Hugo, Helmut Tervooren eds. 36., neugestaltete und erweiterte Auflage I Texte, Stuttgart: S. Hirzel Verlag, 1977.

[320] Mošin, Vladimir A. *Anchor Watermarks*. Amsterdam: Paper Publications Society (Labarre Foundation), 1973. Monumenta Chartæ Papyraceæ Historiam Illustrantia; v. 13.

[321] Mréjen, Valérie, *Eau sauvage*, 2004.

[322] 'Zuigan calls himself "Master"'. Mumon Ekai. (In *The Gateless Gate*, Case 12.)

[323] Mustafa Siddik Altinakar, René Walther (éds.), *Hydraulique fluviale. Tome 16, Écoulement et phénomènes de transport dans les canaux à géométrie simple*, 2008.

[324] Mégnin, Michel (éd.), *Tunis 1900 Lehnert & Landrock photographes*,2005.

[325] Nabokov, Vladimir *Pnin* (1953) , p.14

[326] *The New Penguin English Dictionary*. London: Penguin Books (1986).

[327] Derived from New Zealand Parliament, Legislative Council. *Nga Korero Paramete: 1881-1885* (*New Zealand Electronic Text Collection*). 2008 Wellington, New Zealand. https://nzetc.victoria.ac.nz/tei-source/NZPaV01NgaK.xml.

[328] *Njal's saga*. tr. Magnus Magnusson and Hermann Palsson. Penguin. (1960), chapter 12, p.60.

[329] O'Casey, Sean. *Time to go* (1951).

[330] Ollagnier, Jeanne, *Main*, 2008.

[331] Herodotus. *On Libya*, from the *Histories*.

[332] *[Letter of Capt. E. Hopkins. Providence, 10 Sep 1764]*

[333] "[I am a curse]" in *Original Manuscript Drafts for the Song of Myself*, Collection MSS3829, part of the Papers of Walt Whitman, Clifton Waller Barrett Library of American Literature, Albert and Shirley Small Special Collections Library, University of Virginia.

[334] Ormesson, Jean d', *La Douane de mer*, 1993.

[335] Orwell, George. *Nineteen-Eighty-Four*. London: Gollancz (1949).

[336] 大隅亜希子. 1996. '律令制下における権衡普及の実態: 海産物の貢納単位を中心として'. 史論 49 . pp. 22–44. http://id.nii.ac.jp/1632/00015761/.

[337] Owen, Wilfred. *Dulce et decorum est*, from autograph manuscript in the English Faculty Library, Oxford University.

[338] Owenson, Sydney. *The Wild Irish Girl*. Ed. Kathryn Kirkpatrick. Oxford: Oxford University Press, 1999, p. 17.

[339] 擺夷文經，中央研究院歷史語言研究所。

[340] Pascal, Blaise, 'Le Mémorial', in *Pensées*,1670.

[341] Payne, J. 'Report on the compatibility of J P French's spoken corpus transcription conventions with the TEI guidelines for transcription of spoken texts', *Working Paper*, Dec 1992, NERC WP8/WP4 122.

[342] Peacock, Thomas Love. *Gryll Grange* (1861), chapter 1.

[343] Pennac, Daniel , *La Petite marchande de prose*, 1978.

[344] Perec, Georges, *Espèces d'espaces*, 1974.

[345] Perec, Georges, *La Vie mode d'emploi : romans*, 1978.

[346] Perec, Georges , *Les Choses*, 1965.

[347] 'Partial family tree for Bertrand Russell' based on an example in Pereira, Fernando C.N. and Stuart M. Shieber, *Prolog and Natural Language Analysis*, Stanford: Center for the Study of Language and Information. (1987), p.22.

[348] Pernoud, Régine, *La Femme au temps des Cathédrales*, 1982.

[349] Perreault, Simon, *vCard Format Specification* Internet Engineering Task Force (IETF) (2011). https://datatracker.ietf.org/doc/html/rfc6350.

[350] *Petit Larousse en Couleurs.* Paris: Larousse, (1990).

[351] Philipott, Thomas *Poems.* (1646).

[352] Pinsky, Robert. 'Essays on Psychiatrists' in *Sadness and Happiness* (1975).

[353] Pix, Mary. *The False Friend* (1699).

[354] Plautus, Titus Macchius. *Asinaria, or the Comedy of Asses.* Edited and translated by Wolfgang de Melo. Loeb Classical Library 60. Cambridge, MA: Harvard University Press (2011).

[355] Plautus, Titus Macchius. *Menaechmi.*

[356] Polac, Michel, 'Un fils unique : script', *l'Avant-scène cinéma*, n° 99, janvier. 1970.

[357] Pope, Alexander. 'The Rape of the Lock' (1714) III.7.

[358] Pope, Alexander. *An Essay on Criticism* (1711).

[359] Pope, Alexander. *Dunciad Variorum* (1729), III.284.

[360] From letter 'JK' found in *Poulson's Daily Advertiser,* 8 Oct 1835.

[361] Proust, Marcel, *A la recherche du temps perdu : Du côté de chez Swann*, 1913.

[362] Queneau, Raymond. *Exercices de style.* Paris: Gallimard, (1947), p. 192.

[363] 瓊瑤，《還珠格格》。

[364] Queneau, Raymond , *Journaux*, 1914-1965.

[365] Rabelais, François. *Gargantua* Lyon, 1542

[366] Rabelais, François *Tiers livre des faictz et dictz Heroïques du noble Pantagruel...* (1546), adapted from the website at http://www.bvh.univ-tours.fr/Epistemon/

[367] Renard, Jules, *Journal : 1887-1910*, 1910.

[368] Reference from the bibliography in Reps, Thomas William and Teitelbaum, Tim eds. *The Synthesizer Generator: A system for constructing language-based editors*, New York and Berlin: Springer-Verlag, 1989, p.304.

[369] 日本法隆寺貝葉心經 (Wikipedia Sept 2008).

[370] Richardson, Samuel. *Clarissa; or the History of a Young Lady* (1748), 2 Letter XIV.

[371] Robert, Paul. *Le Petit Robert.* Paris: Dictionnaires Le Robert (1967).

[372] Rohmer, Eric, 'Ma nuit chez Maud : script', *l'Avant-scène cinéma*, n°98, 2001.

[373] Romains, Jules, *Knock*, 1923.

[374] Rostand, Edmond, *Cyrano de Bergerac*, 1898.

[375] Roubaud, Jacques, *La Boucle*, 1993.

[376] Rowling, J. K. 'The Sorting Hat'. In *Harry Potter and the Sorcerer's Stone*, New York: Scholastic, Inc. (1999), chapter 7 , p. 121.

[377] *The Saga of the Volsungs: the Norse epic of Sigurd the Dragon Slayer.* trans. Jesse L. Byock. University of California Press (1990).

[378] *Le Sanctoral du lectionnaire de l'office dominicain (ms. Rome, Sainte-Sabine XIV L 1, f.142r et 189r-230v, Ecclesiasticum officium secundum ordinem fratrum praedicatorum)*, édition électronique par l'Ecole nationale des chartes, d'après l'édition d'Anne-Élisabeth Urfels-Capot (Paris : Ecole nationale des chartes, 2007).

[379] 崎山理. 1985 'インドネシア語'. 講座日本語学 11. 外国語との対照II. 東京:明治書院 61-80

[380] 三毛，〈沙漠中的飯店〉，《撒哈拉的故事》。

[381] 《三家詩鈔》，華培昌，藏於國家圖書館。

[382] Sapir, Edward. *Language: an introduction to the study of speech*, New York: Harcourt, Brace and World, 1921, p.79.

[383] Shakespeare William, *Richard III*, 1597.

[384] Shakespeare, William. *Henry V* In *Mr. William Shakespeares Comedies, Histories, & Tragedies*, London: Jaggard and Blount, 1623.

[385] Shakespeare, William. *Henry V* In *The Works of Shakespeare in seven volumes*, ed. Lewis Theobald, London: Bettesworth, Hitch, Tonson et al, (1733).

[386] Shakespeare, William. *Antony and Cleopatra*, IV.4, 14-21.

[387] Shakespeare, William. *Merchant of Venice*, I.ii, speech 5 (Portia).

[388] Shakespeare, William. *Macbeth*, Act V, Scene 1.

[389] Shakespeare, William. *Antony and Cleopatra* (1623), V. ii.

[390] Shakespeare, William. 'Hamlet' In Stanley Wells and Gary Taylor eds. *The Complete Works*, Oxford: Clarendon Press, 1986, I.i.

[391] Shakespeare, William. *Hamlet*, London:Valentine Simmes. (1603), I.i.

[392] Shakespeare, William. 'The Tempest' In *Mr. William Shakespeares Comedies, Histories, & Tragedies*, London: Jaggard and Blount, (1623).

[393] Shakespeare, William. 'Twelfth Night, or What you Will' In *Mr. William Shakespeares Comedies, Histories, & Tragedies*, London: Jaggard and Blount, (1623).

[394] Shakespeare, William, *The Sonnets* (1609), 18.

[395] Shakespeare, William, *The Sonnets* (1609), 130.

[396] 莎士比亞，《終成眷屬》。

[397] 莎士比亞，《馬克白》。

[398] *Sharing Ancient Wisdoms*, 2013, available: `http://www.ancientwisdoms.ac.uk/`.

[399] Shaw, George Bernard. *Heartbreak House: a fantasia in the Russian manner on English themes* (1919).

[400] Shaw, George Bernard. *Pygmalion*, 1913.

[401] Shields, David. *Dead Languages*, HarperCollins Canada/Perennial Rack, rpt. 1990, p.10.

[402] Sinclair, John ed. *Collins COBUILD English Language Dictionary*. London and Glasgow: Collins, (1987), p. 337 s.v. croissant.

[403] 京劇《四郎探母》，1947年梅蘭芳演出。

[404] Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations*, London. (1776), index to vol. 1.

[405] Smith, Sydney. *Autograph letter*. In Pierpont Morgan library; **?** 11.

[406] Smith, William. *A New Classical Dictionary of Greek and Roman Biography, Mythology, and Geography, Partly Based Upon the Dictionary of Greek and Roman Biography and Mythology*, New York: Harper & Brothers, Publishers, 1860, 1026.

[407] Southey, Robert. Autograph manuscript of *The Life of Cowper*. In Pierpont Morgan MA 412 (**?** 15).

[408] Soyinka, Akinwande Oluwole Wole. *Madmen and Specialists*, London: Methuen (1971).

[409] Spenser, Edmund. *The Faerie Queene: Disposed into twelue bookes, Fashioning XII. Morall vertues.* (1596).

[410] Sterne, Laurence. *The Life and Opinions of Tristram Shandy, Gentleman.* (1760).

[411] Stow, John *A survey of the cities of London and Westminster: containing the original, antiquity, increase, modern estate and government of those cities. Written at first in the year MDXCVIII. By John Stow, citizen and native of London. ... Now lastly, corrected, improved, and very much enlarged: ... by John Strype, ... In six books. ...* (London c.1525 - London 1605).

[412] Sudetic, Chuck *Serbs Tighten Grip On Eastern Bosnia.New York Times,* April 20, 1992.

[413] *SUEtheTrex, Twitter biography.* https://twitter.com/SUEtheTrex. Accessed March 25th, 2020.

[414] Sutherland, L.S. and L.G. Mitchell eds. *The Eighteenth century*, The History of the University of Oxford V, p.178.

[415] 蘇童，《妻妾成群》。

[416] 蘇軾，《定風波》。

[417] Swift, Jonathan. *Travels into Several Remote Nations of the World, in Four Parts. By Lemuel Gulliver...* (1735).

[418] Swift, Jonathan. *Law is a bottomless pit, or the history of John Bull* (1712).

[419] Swinburne, Algernon Charles. *Poems and Ballads (First Series).* London: Chatto & Windus. (1904).

[420] Swinnerton, Frank Arthur. *The Georgian Literary Scene 1910-1935*, 1938, London: J. M. Dent, p. 195.

[421] 台灣，《結婚入盟誓》，1997年5月。

[422] 台灣區三碼與三加二碼郵遞區號

[423] 《台灣叢書》，伊能嘉矩著，藏於國立台灣大學圖書館。(http://catalog.ndap.org.tw/?URN=2155366 (Aug 2008)).

[424] 唐．白居易．琵琶行

[425] Tasovac, T., Romary, L., Banski, P., Bowers, J., de Does, J., Depuydt, K., Erjavec, T., Geyken, A., Herold, A., Hildenbrandt, V., Khemakhem, M., Lehečka, B., Petrović, S., Salgado, A. and Witt, A. *TEI Lex-0: A baseline encoding for lexicographic data.. Version 0.9.3.* DARIAH Working Group on Lexical Resources., 2018 https://dariah-eric.github.io/lexicalresources/pages/TEILex0/TEILex0.html.

[426] Taylor, John *The Cold Tearme or, The Frozen age, Or, The Metamorphosis of the Riuer of Thames.* London, 1621. [STC (2nd ed.) 23910].

[427] TEI Consortium *Call for Papers - TEI 2022* Newcastle, 2022. https://web.archive.org/web/20220516140643/https://conferences.ncl.ac.uk/tei2022/cfp/

[428] 'The Castle of the Fly', in *Russian Fairy Tales,* translated by Norbert Guterman from the collections of Aleksandr Afanas'ev, illustrations by Alexander Alexeieff, folkloristic commentary by Roman Jakobson (New York: Pantheon Books, 1947, rpt. [n.d.]), p. 25.

[429] *The Daily Telegraph*, 21 Dec 1992.

[430] *The Guardian,* 26 Oct 1992, p. 2

[431] *The Guardian,* 21 Dec 1992, p. 2.

[432] *The Holy Bible, conteyning the Old Testament and the new... appointed to be read in Churches.* (1611), Genesis 1:1.

[433] *The Independent,* 26 Oct 1775, headline.

[434] Thurber, James. *The 13 Clocks* (1950).

[435] Tolkien, J. R. R. *The Monsters and the Critics.* London: George Allen & Unwin (1983).

[436] Townsend, Sue. *The growing pains of Adrian Mole* (1984), p.43.

[437] Trollope, Anthony. *An Autobiography* (1883).

[438] Trollope, Anthony. *North America* (1862).

[439] Truffaut, François et Gruault, Jean, 'Les deux Anglaises et le continent : script', *l'Avant-scène cinéma*, n° 121, nov. 1971.

[440] Tufte, Edward R., *Envisioning Information*, Cheshire: Graphics Press. (1990).

[441] *Ukraine Relief Fund* UBS https://web.archive.org/web/20220307150159/https://www.ubs.com/global/en/ubs-society/philanthropy/optimus-foundation/ukrainerelief.html.

[442] *United States Code* Title 17, Section 107, found at http://www.copyright.gov/title17/92chap1.html#107.

[443] United States District Court for the Middle District of Pennsylvania. *Kitzmiller v. Dover Area School District et al.*: 2005. 04cv2688, p. 33. Available from http://ncse.com/files/pub/legal/kitzmiller/highlights/2005-12-20_Kitzmiller_decision.pdf and transcribed at http://en.wikisource.org/wiki/Kitzmiller_v._Dover_Area_School_District_et_al..

[444] Valéry, Paul, *Monsieur Teste*, 1929.

[445] Vergil (Publius Vergilius Naso). *Aeneid*, I.1.

[446] Verne, Jules, *Cinq semaines en ballon*, 1863.

[447] Verne, Jules, *La Chasse au météore*, 1908.

[448] Verne, Jules,*Vingt mille lieues sous les mers*, 1870.

[449] Vinge, Vernor *Across realtime* ch 10 (1986).

[450] Viton de Saint-Allais, Nicolas , *Dictionnaire encyclopédique de la noblesse de France*, 1816.

[451] *Vóluspá* recto of folio 5 of the unique manuscript of the Elder Edda. Codex Regius, ed. L. F. A. Wimmer and F. Jónsson (Copenhagen 1891).

[452] Wanklyn, M.D.G. et al. *Gloucester Port Books, 1575-1765*. Available from http://discover.ukdataservice.ac.uk/catalogue?sn=3218.

[453] Wanton, Joseph. *Unpublished letter to Nicholas Brown and Co*, 1761 Brown University Steering Committee on Slavery and Justice: Repository of Historical Documents. (http://library.brown.edu/cds/slaveryandjustice/).

[454] Warriner, John E. *English Composition and Grammar* (1988), p.280.

[455] *Webster's Seventh Collegiate Dictionary*. Springfield, Mass. G. & C. Merriam Co. (1975).

[456] 魏飴，《小說鑑賞入門》，台北：萬卷樓，1999。

[457] Weil, Simone, 'Lettres à Antonio Atarès', in *Oeuvres complètes*1988-.

[458] Williams, Nigel. *The Wimbledon Poisoner* (1990), p. 204.

[459] Winock, Michel, *Jeanne et les siens : récit*, 2003.

[460] Wölfflin, Heinrich, trans. Hottinger, Marie Donald Mackie (1932). *Principles of art history: the problem of the development of style in later art.*. Originally published as *Kunstgeschichtliche Grundbegriffe* (1915).

[461] Woolf, Virginia. *Mrs Dalloway* (1925), p.64, p.65.

[462] Wordsworth, William. 'Scorn not the sonnet' in *Poetical Works* (1827).

[463] Wordsworth, William. *The Prelude* (1850).

[464] Wrenn C. L. ed. *Beowulf: with the Finnesburg fragment*, London: Harrap (1953).

[465] 吳承恩，《西遊記》。

[466] Wycherley, William. *The Country Wife* (1675).

[467] 夏宇，〈甜蜜的復仇〉，《備忘錄》。

[468] 蕭紅，《呼蘭河傳》。

[469] 幼莘貢俚賀曉濤之新婚賀詞

[470] 余秋雨

[471] 元曲。馬致遠，《天淨沙：秋思》。

[472] 張錯，《西洋文學術與手冊》，台北：書林，2005。頁201。

[473] 珍．奧斯丁，《傲慢與偏見》。

[474] 中國青年守則

[475] 中國學位論文全文數據庫

[476] 周慧玲，《表演中國：女明星表演文化視覺政治1910-1945》，2004。

[477] 朱自清，《憶》跋。

[478] Zimman, Lal. *Lal Zimman* http://lalzimman.com/bio.html, accessed February 1, 2021.

[479] Zupko, Ronald Edward *British Weights & Measures: A History from Antiquity to the Seventeenth Century*. Madison: University of Wisconsin Press, 141-151.

[480] No source, made up for these *Guidelines*.

[481] Example is copied from the source of these *Guidelines*.

[482] Undetermined.

## F.2 Works Cited Elsewhere in the Text of these Guidelines

[1] Scott E. Fahlman *"Joke" Conversation Thread in which the :-) Was Invented* http://www.cs.cmu.edu/~sef/Orig-Smiley.htm Donald E. Knuth. *Literate Programming*, CSLI Lecture Notes 27 0-937073-80-6. 1992. Stanford, California: Center for the Study of Language and Information. Renato G. Mazzolini. 'Colonialism and the Emergence of Racial Theories'. 10.1017/9781107705647.032. *Reproduction: Antiquity to the Present Day*, Nick Hopwood, Rebecca Flemming, Lauren Kassell (eds.) 2018. Cambridge: Cambridge University Press. pp. 361-374. Gayle Rubin. 'The Traffic in Women: Notes on the 'Political Economy' of Sex'. https://philpapers.org/archive/RUBtti.pdf *Toward an Anthropology of Women*, Rayna R. Reiter (ed.) 1975. New York: Monthly Review Press. pp. 157–210, 165. A. G. Petty. *English literary hands from Chaucer to Dryden*, 1977. London: Edward Arnold. pp. 22–25. Michael Beißwenger, Thomas Bartz, Angelika Storrer, Swantje Westpfahl. *Tagset and guidelines for the PoS tagging of language data from genres of computer-mediated communication / social media*, https://nbn-resolving.org/urn:nbn:de:bsz:mh39-50650 2015-09-13. Michael Beißwenger, Harald Lüngen. *CMC-core: a schema for the representation of CMC corpora in TEI*, https://journals.openedition.org/corpus/4553 *Corpus 20 (Special issue "Traitements, standardisation et analyse des corpus de communication médiée par les réseaux sociaux)* Céline Poudat, Ciara R. Wigham, Loïc Liégeois (eds.) 2020. Gerry and demolog. *Blog comments to "Scheinzwerge oder Viele Probleme werden größer, wenn man sie anpackt (Griechenland)"*, https://scilogs.spektrum.de/wild-dueck-blog/scheinzwerge-oder-viele-probleme-werden-groesser-wenn-man-sie-anpackt-griechenland/ 29-30 July 2015. *WILD DUECK BLOG*, Gunter Dück. https://scilogs.spektrum.de/wild-dueck-blog/. SciLogs

[2] Jossi et al. (2006-): *Talk:Eiffel (programming language)/Archive_1.* URL: https://en.wikipedia.org/wiki/Talk:Eiffel_(programming_language)/Archive_1. English Wikipedia talk page, Wikimedia Foundation.

[3] MiszaBot I et al. (2011-): *Talk:Astronomical object.* URL: https://en.wikipedia.org/wiki/Talk:Astronomical_object. English Wikipedia talk page, Wikimedia Foundation.

[4] OnkelSchuppig et al. (2001-): *Diskussion:FKM-Richtlinie.* URL: https://de.wikipedia.org/wiki/Diskussion:FKM-Richtlinie. German Wikipedia talk page, Wikimedia Foundation. *Mobile Communication Database 2 (MoCoDa2)*, Michael Beißwenger, Evelyn Ziegler, Marcel Fladrich, Wolfgang Imo, Katharina König (eds.) visited 30 March 2022. https://db.mocoda2.de/c/home. *Dortmund Chat Corpus*, Angelika Storrer, Michael Beißwenger (eds.) 2017. http://hdl.handle.net/10932/00-03B0-14FA-A8D0-0F01-F. Chanier Thierry, Poudat Céline, Sagot Benoit, Antoniadis Georges, Wigham Ciara R., Hriba Linda, Longhi Julien, Seddah Djamé. *The CoMeRe corpus for French: structuring and annotating heterogeneous CMC genres*, *JLCL (Journal of Language Technology and Computational Linguistics) (Special issue on « Building And Annotating Corpora Of Computer-Mediated Discourse: Issues and Challenges at the Interface of Corpus and Computational Linguistics)*, 2014. ??ttp://www.jlcl.org/2014_Heft2/Heft2-2014.pdf. 2 pp. 1-31. Louis Alexander Cotgrove (ed.) *Nottinghamer Korpus Deutscher YouTube-Sprache (The NottDeuYTSch Corpus)*, 2018. http://hdl.handle.net/11372/LRT-4806. Ciara Wigham, Thierry Chanier. *Interactions between text chat and audio modalities for L2 communication and feedback in the synthetic world Second Life*, 10.1080/09588221.2013.851702. *Computer Assisted Language Learning*, 2015. 23 (3) . Lou Burnard. 'Report of Workshop on Text Encoding Guidelines'. *Literary & Linguistic Computing* 1988. 3. Lou Burnard, C. Michael Sperberg-McQueen. 'The Design of the TEI Encoding Scheme'. 10.1007/BF01830314. *Computers and the Humanities* 1995. 29 (1) p. 17–39. (Reprinted in Ide and Veronis (eds.) (1995), pp. 17-40)Lou Burnard, Sebastian Rahtz. *RelaxNG with Son of ODD*, http://www.mulberrytech.com/Extreme/Proceedings/html/2004/Burnard01/EML2004Burnard01.pdf *Proceedings of Extreme Markup Languages 2004*, 2004. M. B. Parkes. *English Cursive Book Hands 1250–1500*, 1969. Oxford: Clarendon Press. *British Literary Manuscripts. Series*

*2: from 1800 to 1914*, Verlyn Klinkenborg, Herbert Cahoon. 1981. New York: Pierpont Morgan Library. Paul Grosso, Eve Maler, Jonathan Marsh, Norman Walsh (eds.) *XPointer Framework*, https://www.w3.org/TR/xptr-framework/ 25 March 2003. W3C. Paul Grosso, Eve Maler, Jonathan Marsh, Norman Walsh (eds.) *XPointer element() Scheme*, https://www.w3.org/TR/xptr-element/ 25 March 2003. W3C. *XHTML™ 1.0 The Extensible HyperText Markup Language (Second Edition)*, https://www.w3.org/TR/xhtml/ 26 January 2000. W3C. Dave Ragget, Arnaud Le Hors, Ian Jacobs (eds.) *HTML 4.01 Specification*, https://www.w3.org/TR/html401/ 24 December 1999. W3C. David Carlisle, Patrick Ion, Robert Miner, Nico Poppelier (eds.) *Mathematical Markup Language (MathML) Version 2.0 (Second edition)*, https://www.w3.org/TR/MathML2/ 21 October 2003. W3C. Paul V. Biron, Ashok Malhotra (eds.) *XML Schema Part 2: Datatypes Second Edition*, https://www.w3.org/TR/2004/REC-xmlschema-2-20041028/ 28 October 2004. W3C. Anders Berglund (ed.) *Extensible Stylesheet Language (XSL) Version 1.1*, https://www.w3.org/TR/xsl11/ 5 December 2006. W3C. James Clark (ed.) *XSL Transformations (XSLT) Version 1.0*, https://www.w3.org/TR/xslt/ 16 November 1999. W3C. Michael Kay (ed.) *XSL Transformations (XSLT) Version 2.0*, https://www.w3.org/TR/xslt20/ 23 January 2007. W3C. Michael Kay (ed.) *XSL Transformations (XSLT) Version 3.0*, https://www.w3.org/TR/xslt-30/ 8 June 2017. W3C. Tim Bray, Jean Paoli, C. M. Sperberg-McQueen, Eve Maler, François Yergau (eds.) *Extensible Markup Language (XML) Version 1.0 (Fourth edition)*, https://www.w3.org/TR/REC-xml/ 16 August 2006. W3C. Bert Bos, Tantek Çelik, Ian Hickson, Håkon Wium Lie (eds.) *Cascading Style Sheets Level 2 Revision 1*, https://www.w3.org/TR/CSS2/ 7 June 2011. W3C. Håkon Wium Lie, Bert Bos (eds.) *Cascading Style Sheets, Level 1*, https://www.w3.org/TR/REC-CSS1/ 11 January 1999. W3C. fantasai, Koji Ishi (eds.) *CSS Writing Modes Level 3 (W3C Candidate Recommendation)*, https://www.w3.org/TR/css-writing-modes-3/ 15 December 2015. W3C. Simon Fraser, Dean Jackson, Edward O'Connor, Dirk Schulze (eds.) *CSS Transforms Module Level 1 (W3C Working Draft)*, https://www.w3.org/TR/css-transforms/ 26 November 2013. W3C. Frank Manola, Eric Miller, Brian McBride (eds.) *RDF 1.1 Primer*, https://www.w3.org/TR/rdf11-primer/ 24 June 2014. W3C. Jonathan Marsh, Richard Tobin (eds.) *XML Base (Second Edition)*, https://www.w3.org/TR/xmlbase/ 28 January 2009. W3C. James Clark, Steve DeRose (eds.) *XML Path Language (XPath) Version 1.0*, https://www.w3.org/TR/xpath/ 16 November 1999. W3C. Anders Berglund, Scot Boag, Mary F. Fernández, Michael Kay, Jonathan Robie, Jérôme Siméon (eds.) *XML Path Language (XPath) 2.0*, https://www.w3.org/TR/xpath20/ 23 January 2007. W3C. Jonathan Robie, Don Chamberlin, Michael Dyck, Jon Snelson (eds.) *XML Path Language (XPath) 3.0*, https://www.w3.org/TR/xpath-30/ 8 April 2014. W3C. Jonathan Robie, Michael Dyck, Josh Spiegel (eds.) *XML Path Language (XPath) 3.1*, https://www.w3.org/TR/xpath-31/ 21 March 2017. W3C. Tim Bray, Dave Hollander, Andrew Laymon, Richard Tobin (eds.) *Namespaces in XML 1.0 (second edition)*, https://www.w3.org/TR/xml-names/ 16 August 2006. W3C. Eric van der Vlist. *RELAX NG*, http://books.xmlschemata.org/relaxng/page2.html 2004. O'Reilly. A. Renear, E. Mylonas, D. Durand. 'Refining our notion of what text really is: the problem of overlapping hierarchies'. Nancy Ide, Susan Hockey (eds.) *Research in Humanities Computing*, 1996. Oxford University Press. Robert Sanderson, Paolo Ciccarese, Benjamin Young (eds.) *Web Annotation Data Model*, https://www.w3.org/TR/annotation-model/ 23 February 2017. W3C. Robert Sanderson, Paolo Ciccarese, Benjamin Young (eds.) *Web Annotation Vocabulary*, https://www.w3.org/TR/annotation-vocab/ 23 February 2017. W3C. *The Unicode Standard, Version 5.0*, https://www.unicode.org/ Unicode Consortium. 2006. Addison-Wesley Professional. Addison Phillips, Mark Davis (eds.) *Tags for Identifying Languages*, RFC 4646. 2006. IETF. Addison Phillips, Mark Davis (eds.) *Matching of Language Tags*, RFC 4647. 2006. IETF. Mark Davis, Ken Whistler, Asmus Freytag. *Unicode Character Database*, https://www.unicode.org/Public/UNIDATA/UCD.html 2006. Unicode Consortium. Fernando C. N. Pereira. *Grammars and logics of partial information*, 1987. Menlo Park, CA: SRI International. SRI International Technical Note 420. Bob Carpenter. *The logic of typed feature structures*, 1992. Cambridge: Cambridge University Press. Cambridge Tracts in Theoretical Computer Science 32. Stuart Shieber. *An Introduction to Unification-based Approaches to Grammar*, CSLI Lecture Notes 4. 1986. Palo Alto, CA: Center for the Study of Language and Information. *Anglo-American Cataloguing Rules*, Second Edition, 2002 revision, 2005 update. 2002–2005. Chicago, Ottawa: American Library Association. Canadian Library Association. John Burrows. *Computation into Criticism: A Study of Jane Austen's Novel and an Experiment in Method*, 1987. Oxford: Clarendon Press. Klaus Mattheier, Ulrich Ammon, Peter Trudgill (eds.) *Sociolinguistics, Soziolinguistik, An international handbook of the science of language and society, Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*, 1988. Berlin, New York: De Gruyter. I pp. 271 and 274. J. A. Edwards, M. D. Lampert (eds.) *Talking Language: Transcription and Coding of Spoken Discourse*, 1993. Hillsdale, N.J.: Lawrence

Erlbaum Associates. Asmus Freytag. *The Unicode Character Property Model*, Unicode Technical Report #23 https://www.unicode.org/reports/tr23/ 2006. Mark Davis, Aharon Lanin, Andrew Glass. *Unicode Bidirectional Algorithm*, Unicode Standard Annex #9 https://www.unicode.org/reports/tr9/ 2017-05-04. r. 37. Stig Johansson. 'Encoding a Corpus in Machine-Readable Form'. Sue Atkins, Antonio Zampolli (eds.) *Computational Approaches to the Lexicon: An Overview*, 1994. Oxford: Oxford University Press. Stig Johansson, Lou Burnard, Jane Edwards, And Rosta. *Working Paper on Spoken Texts*, *TEI document TEI AI2 W1*, 1991. Brian MacWhinney. *CHAT Manual*, 1988. Pittsburgh: Dept of Psychology, Carnegie-Mellon University. pp. 87ff. Bengt Loman, Nils Jørgensen. *Manual for analys och beskrivning av makrosyntagmer*, 1971. Lund: Studentlitteratur. Robert A. Amsler, Frank W. Tompa. 'An SGML-Based Standard for English Monolingual Dictionaries'. *Information in Text*, *Fourth Annual Conference of the U[niversity of] W[aterloo] Centre for the New Oxford English Dictionary*, (Fourth Annual Conference of the U[niversity of] W[aterloo] Centre for the New Oxford English Dictionary, October 26-28, 1988, Waterloo, Canada) October 1988. Waterloo, Canada. pp. 61-79. N. Calzolari, C. Peters, A. Roventini. *Computational Model of the Dictionary Entry: Preliminary Report*, *Acquilex: Esprit Basic Research Action No. 3030, Six-Month Deliverable*, April 1990. Pisa. John Fought, Carol Van Ess-Dykema. *Toward an SGML Document Type Definition for Bilingual Dictionaries*, *TEI working paper TEI AIW20*, available from the TEI.. Nancy Ide, Jean Veronis. 'Encoding Print Dictionaries'. *Computers and the Humanities* 1995. 29 pp. 167-195. Nancy Ide, Jacques Le Maitre, Jean Veronis. 'Outline of a Model for Lexical Databases'. *Information Processing and Management* 1993. 29 (2) pp. 159-186. Nancy Ide, Jean Veronis, Susan Warwick-Amstrong, Nicoletta Calzolari. 'Principles for Encoding machine readable dictionaries'. *Proceedings of the Fifth EURALEX International Congress, EURALEX'92*, (Fifth EURALEX International Congress, EURALEX'92, University of Tampere, Finland) 1992. The DANLEX Group. 'Descriptive tools for electronic processing of dictionary data'. *Lexicographica, Series Maior* 1987. Tübingen: Niemeyer. Agnès Tutin, Jean Veronis. 'Electronic dictionary encoding: customizing the TEI Guidelines'. *Proceedings of the Eighth Euralex International Congress*, (Eighth Euralex International Congress) 1998. N. Ide, A. Kilgarriff, L. Romary. 'A Formal Model of Dictionary Structure and Content'. *Proceedings of Euralex 2000*, (Euralex 2000) 2000. Stuttgart. pp. 113-126. William A. Gale, Kenneth W. Church. 'Program for aligning sentences in bilingual corpora'. *Computational Linguistics* 1993. 19 pp. 75-102. G. N. Leech, R. G. Garside. 'Running a Grammar Factory'. S. Johansson, A.-B. Stenstrøm (eds.) *English Computer Corpora: Selected Papers and Research Guide*, 1991. Berlin, New York: de Gruyter. Mouton. pp. pp. 15-32.. I. Marshall. 'Choice of Grammatical Word Class without Global Syntactic Analysis: Tagging Words in the LOB Corpus'. *Computers and the Humanities* 1983. 17 pp. 139-50. R. G. Garside, G. N. Leech, G. R. Sampson. *The Computational Analysis of English: a Corpus-Based Approach*, 1991. Oxford: Oxford University Press. Gary Chartrand, Linda Lesniak. *Graphs and Digraphs*, 1986. Menlo Park, CA: Wadsworth. R. Jackendoff. 'X-Bar Syntax: A study of phrase structure'. *Linguistic Inquiry Monograph* 1977. 2. M. Kytö, M. Rissanen. 'The Helsinki Corpus of English Texts'. M. Kytö, O. Ihalainen, M. Rissanen (eds.) *Corpus Linguistics: hard and soft*, 1988. Amsterdam: Rodopi. Steven DeRose. 'Markup overlap: a review and a horse'. http://www.mulberrytech.com/Extreme/Proceedings/html/2004/DeRose01/EML2004DeRose01.html *Proceedings of Extreme Markup Languages 2004*, 2004. Andreas Witt. *Multiple Informationsstrukturierung mit Auszeichnungssprachen. XML-basierte Methoden und deren Nutzen für die Sprachtechnologie*, 2002. (Ph D thesis, Bielefeld University) (See also http://xml.coverpages.org/Witt-allc2002.html)Mirco Hilbert, Oliver Schonefeld, Andreas Witt. 'Making CONCUR work'. http://www.mulberrytech.com/Extreme/Proceedings/html/2005/Witt01/EML2005Witt01.xml *Proceedings of Extreme Markup Languages 2005*, 2005. Alex Dekhtyar, Ionut E. Iacob. *A framework for management of concurrent XML markup*, http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1104&context=csse_fac 2005. H. V. Jagadish, Laks V. S. Lakshmanan, Monica Scannapieco, Divesh Srivastava, Nuwee Wiwatwattana. *Colorful XML: one hierarchy isn't enough*, http://web.archive.org/web/20150706093254/http://www2.research.att.com/~divesh/papers/jlssw2004-mct.pdf 2004. Noureddine Chatti, Suha Kaouk, Sylvie Calabretto, Jean Marie Pinon. 'MultiX: an XML based formalism to encode multistructured documents'. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.1525&rep=rep1&type=pdf *Proceedings of Extreme Markup Languages 2007*, 2007. Patrick Durusau, Matthew Brook O'Donnell. 'Coming down from the trees: next step in the evolution of markup?'. *Proceedings of Extreme Markup Languages 2002*, 2002. Jeni Tennison, Wendell Piez. 'The layered markup and annotation language'. *Proceedings of Extreme Markup Languages Conference*, 2002. Claus Huitfeldt, C. Michael Sperberg-McQueen. *TexMECS: An experimental markup meta-language for complex documents*, http://mlcd.blackmesatech.com/mlcd/2003/Papers/texmecs.html 2001. D. Terence Langendoen, Gary F. Simons. 'A rationale

for the TEI recommendations for feature-structure markup,'. *Computers and the Humanities* 1995. 29 pp. 167-195. British Standards Institute. *BS 5605:1990: Recommendations for Citing and Referencing Published Material*, 1990. British Standards Institute. *BS 6371:1983: Recommendations for Citation of Unpublished Documents*, 1983. Deutsches Institut für Normung. *DIN 1505-2: Titelangaben von Dokumenten; Zitierregeln*, 1984. Die Deutsche Bibliothek. *Regeln für die alphabetische Katalogisierung in wissenschaftlichen Bibliotheken RAK-WB*, 2006. International Organization for Standardization. *ISO 690:1987: Information and documentation – Bibliographic references – Content, form and structure*, 1987. International Organization for Standardization. *ISO 8601:2004: Data elements and interchange formats — Information interchange — Representation of dates and times*, 2004. International Organization for Standardization. *ISO 12620:2009: Terminology and other language and content resources – Specification of data categories and management of a Data Category Registry for language resources*, http://www.iso.org/iso/catalogue_detail?csnumber=37243 2009. International Organization for Standardization. *ISO 19136:2007: Geographic information — Geography Markup Language (GML)*, 2006. International Organization for Standardization. *ISO/IEC 19757-3:2006: Information technology — Document Schema Definition Languages (DSDL) – Part 3: Rule-based validation – Schematron*, 2006. Istituto Centrale per il Catalogo Unico. *Regole italiane di catalogazione per autori*, 1979. National Information Standards Organization. *ANSI/NISO Z39.29 – 2005 (R2010) Bibliographic References*, 2010. *ISBD: International Standard Bibliographic Description*, 2011. Berlin, München, De Gruyter Saur. *IFLA Series on Bibliographic Control*, 44. Федеральное агентство по техническому регулированию и метрологии (РОССТАНДАРТ). *ГОСТ Р 7.0.5-2008: Система стандартов по информации, библиотечному и издательскому делу. Библиографическая ссылка. Общие требования и правила составления*, 2008. Федеральное агентство по техническому регулированию и метрологии (РОССТАНДАРТ). *ГОСТ 7.1—2003. Система стандартов по информации, библиотечному и издательскому делу. Библиографическая запись. Библоиграфическое описание. Общие требования и правила составления*, 2003. DCMI Usage Board. *Dublin Core™ Metadata Element Set, Version 1.1: Reference Description*, https://www.dublincore.org/specifications/dublin-core/dces/ 2012-06-14.

## F.3 Reading List

The following lists of readings in markup theory and the TEI derive from work originally prepared by Susan Schreibman and Kevin Hawkins for the TEI Education Special Interest Group, recoded in TEI P5 by Sabine Krott and Eva Radermacher. They should be regarded only as a snapshot of work in progress, to which further contributions and corrections are welcomed (see further http://www.tei-c.org/Support/Learn/tei_bibliography.xml).

### Theory of Markup and XML

[Barnard et al. (1995)] David T. Barnard, Lou Burnard, Jean-Pierre Gaspart, Lynne A. Price, C. Michael Sperberg-McQueen, Giovanni Battista Varile. 'Hierarchical Encoding of Text: Technical Problems and SGML Solutions'. 10.1007/BF01830617. http://www.tei-c.org/Vault/ML/mlw18.ps *Computers and the Humanities* 1995. 29 (3) p. 211–231.

[Barnard et al. (1996)] David T. Barnard, Lou Burnard, C. Michael Sperberg-McQueen. 'Lessons learned from using SGML in the Text Encoding Initiative'. 10.1016/0920-5489(95)00035-6. *Computer Standards & Interfaces* 1996. 18 (1) p. 3–10.

[Burnard (1991)] Lou Burnard. 'What is SGML and how does it help?'. http://www.tei-c.org/Vault/ED/EDW25/ Daniel Greenstein (ed.) *Modelling Historical Data: Towards a Standard for Encoding and Exchanging Machine-readable Texts*, 1991. St Katherinen: Max-Planck-Institut für Geschichte In Kommission bei Scripta Mercaturae Verlag. p. 81–91. Halbgraue Reihe zur Historischen Fachinformatik Herausg. von Manfred Thaller serie A 11. (Revised version published as Burnard (1995))

[Burnard (1995)] Lou Burnard. 'SGML on the Web: Too Little Too Soon or Too Much Too Late?'. http://users.ox.ac.uk/~lou/Belux/ *Computers & Texts* 1995. 15 p. 12–15.

[Burnard (1995)] Lou Burnard. 'What is SGML and How Does It Help?'. 10.1007/BF01830315. http://xml.coverpages.org/burnardw25-index.html *Computers and the Humanities* 1995. 29 (1) p. 41–50. (Reprinted in Ide and Veronis (eds.) (1995), pp. 41-50)

[Burnard (1999)] Lou Burnard. *Is Humanities Computing an Academic Discipline? or, Why Humanities Computing Matters*, http://www.iath.virginia.edu/hcs/burnard.html http://www.iath.

virginia.edu/hcs/ 1999. (Presented at an interdisciplinary seminar at the Institute for Advanced Technology in the Humanities, University of Virginia, November 1999.)

[Burnard (1999)] Lou Burnard. 'Using SGML for Linguistic Analysis: The Case of the BNC'. http://users.ox.ac.uk/~lou/papers/sgml96.sgm *Markup Languages Theory and Practice* 1999. Cambridge, Massachusettes: MIT Press. 2 p. 31–51. (Also published in Moser et al. (eds.) (2001), pp. 53–72)

[Burnard et al. (1999)] Lou Burnard, Elizabeth Lalou, Peter Robinson. 'Vers un Standard Européen de Description des Manuscrits: Le Projet Master'. *Documents Numeriques Les Documents Anciens* 1999. Paris: Hermes Science Publications. 3 (1–2) pp. 151-169.

[Burnard (1999)] Lou Burnard. *XML: The Dream and the Reality*, http://users.ox.ac.uk/~lou/papers/euro99.xml 1999. (Closing plenary address at the XML Europe Conference, Granada, May 1999)

[Burnard et al. (2000)] Lou Burnard, Claudia Claridge, Josef Schmied, Rainer Siemund. 'Encoding the Lampeter Corpus'. http://users.ox.ac.uk/~lou/papers/glasgie.xml *DRH98: Selected Papers from Digital Resources for the Humanities*, 2000. London: Office for Humanities Communication.

[Burnard (2000)] Lou Burnard. *From Two Cultures to Digital Culture: The Rise of the Digital Demotic*, http://users.ox.ac.uk/~lou/wip/twocults.html 2000. (Presented at CLIP, Alicante) (Published in Italian as Burnard (2001))

[Burnard (2001)] Lou Burnard. 'Dalle «Due Culture» Alla Cultura Digitale: La Nascita del Demotico Digitale'. Translated by Federico Pellizi*Il Verri Nella Rete* 2001. Milano: Monogramma. 16 p. 9–22.

[Burnard (2001)] Lou Burnard. 'On the Hermeneutic Implications of Text Encoding'. http://users.ox.ac.uk/~lou/wip/herman.htm Domenico Fiormonte, Jonathan Usher (eds.) *New Media and the Humanities: Research and Applications*, 2001. Oxford: Humanities Computing Unit. p. 31–38.

[Burnard (2005)] Lou Burnard. 'Encoding Standards for the Electronic Edition'. http://nl.ijs.si/e-zrc/bib/eziss-Burnard.pdf Matija Ogrin (ed.) *Znanstvene Izdaje in Elektronski Medij*, *Scholarly Editions and the Digital Medium*, 2005. Ljubljana: Studia Litteraria ZRC ZAZU. p. 12–67.

[Burnard (2005)] Lou Burnard. 'Metadata for corpus work'. http://users.ox.ac.uk/~lou/wip/metadata.html Martin Wynne (ed.) *Developing Linguistic Corpora: A Guide to Good Practice*, 2005. Oxford: Oxbow Books. p. 30–46.

[Burnard et al. (eds.) (2006)] Lou Burnard, Katherine O'Brien O'Keefe, John Unsworth (eds.) *Electronic Textual Editing*, http://www.tei-c.org/Vault/ETE/ 2006. New York: Modern Languages Association.

[Buzzetti (2002)] Dino Buzzetti. 'Digital Representation and the Text Model'. http://muse.jhu.edu/journals/new_literary_history/v033/33.1buzzetti.html *New Literary History* 2002. 33 (1) p. 61–88.

[Caton (2001)] Paul Caton. 'Markup's Current Imbalance'. *Markup Languages: Theory and Practice* 2001. 3 (1) p. 1–13. (This paper was proceeded by reports at the Joint Annual Conference of the Association for Computers and the Humanities and the Association for Literary and Linguistic Computing in 1999 (Charlottesville, Virginia) and Extreme Markup Languages 2000 (Montreal, Canada))

[Chen and Yu (2003)] Ruey-Shun Chen, Shien-Chiang Yu. 'Developing an XML Framework for Metadata System'. http://dl.acm.org/citation.cfm?id=963653 *Proceedings of the 1st International Symposium on Information and Communication Technologies*, 2003. Dublin. p. 267–272. ACM International Conference Proceeding Series 49. (This paper was presented in a session entitled "Electronic Document Technology.")

[Coombs (1986)] James H. Coombs. *Information Management System for Scholars*, Technical Memorandum TM 69–2. 1986. Providence: Brown Computer Center.

[Coombs et al. (1987)] James H. Coombs, Allen Renear, Steven J. DeRose. 'Markup Systems and The Future of Scholarly Text Processing'. 10.1145/32206.32209. http://xml.coverpages.org/coombs-hallgren.html http://xml.coverpages.org/coombs.html *Communications of the ACM* 1987. 30 (11) p. 933–947. (Reprinted with new commentary in Landow and Delany (eds.) (1993), pp 85–118)

[Cover (2005)] Robin Cover. *Markup Languages and (Non-) Hierarchies*, http://xml.coverpages.org/hierarchies.html 2005. (Technology report from the Cover Pages)

[DeRose et al. (1990)] Steven J. DeRose, David G. Durand, Elli Mylonas, Allen H. Renear. 'What is Text, Really?'. *Journal of Computing in Higher Education* 1990. 1 (2) p. 3–26. (Republished (DeRose et al. (1997)) as a "classic reprint" with invited commentary and authors' replies in the ACM/SIGDOC)

[DeRose (1995)] Steven J. DeRose. *Structured Information: Navigation, Access, and Control*, http://xml.coverpages.org/deroseStructure.html 1995. (Paper presented at the Berkeley Finding Aid Conference, April 4–6, 1995)

[DeRose et al. (1997)] Steven J. DeRose, David G. Durand, Elli Mylonas, Allen H. Renear. 'What is Text, Really?'. 10.1145/264842.264843. *Journal of Computer Documentation* 1997. 21 (3) p. 1–24.

[Goldfarb (1981)] Charles F. Goldfarb. 'A Generalized Approach to Document Markup'. http://users.nyct.net/~aray/notes/igm.html *Proceedings of the ACM SIGPLAN SIGOA Symposium on Text Manipulation*, 1981. New York: ACM. 68–73. (Adapted as "Annex A. Introduction to Generalized Markup" in ISO 8879)

[Graham (1999)] Tony Graham. 'Unicode: What Is It and How Do I Use It?'. *Markup Languages: Theory & Practice* 1999. 1 (4) p. 75.

[Hockey (1996)] Susan Hockey. 'Creating and Using Electronic Editions'. Richard J. Finneran (ed.) *The Literary Text in the Digital Age*, 1996. Ann Arbor, MI: University of Michigan Press. p. 1–22.

[Hockey et al. (1999)] Susan Hockey, Allen Renear, Jerome J. McGann. *What is Text? A Debate on the Philosophical and Epistemological Nature of Text in the Light of Humanities Computing Research*, http://www2.iath.virginia.edu/ach-allc.99/proceedings/hockey-renear2.html 1999. (Panel presented at ACH/ALLC 1999)

[Hockey (2000)] Susan Hockey. *Electronic Texts in the Humanities*, 2000. New York, NY: Oxford University Press.

[Huitfeldt (1994)] Claus Huitfeldt. 'Multi-dimensional Texts in a One-dimensional Medium'. 10.1007/BF01830270. *Computers and the Humanities* 1994. 28 (4/5) p. 235–241.

[Huitfeldt (1994)] Claus Huitfeldt. 'Toward a Machine-Readable Version of Wittgenstein's Nachlaß: Some Editorial Problems'. Hans Gerhard Senger (ed.) *Philosophische Editionen. Erwartungen an sie — Wirkungen durch sie*, 1994. Tübingen: Max Niemeyer Verlag. p. 37–43. Beihefte zu editio 6.

[Ide and Veronis (eds.) (1995)] Nancy Ide, Jean Veronis (eds.) *The Text Encoding Initiative: Background and Contexts*, 1995. Dordrecht, Boston: Kluwer Academic Publisher.

[Lamport (1987)] Leslie Lamport. 'Document Production: Visual or Logical?'. http://research.microsoft.com/en-us/um/people/lamport/pubs/pubs.html#document-production *Notices of the American Mathematical Society* 1987. 34 p. 621–624. (Republished as Lamport (1988))

[Lamport (1988)] Leslie Lamport. 'Document Production: Visual or Logical?'. http://www.tug.org/TUGboat/Articles/tb09-1/tb20lamport.pdf *TUGboat* 1988. 9 (1) pp. 8-10.

[Landow and Delany (eds.) (1993)] George P. Landow, Paul Delany (eds.) *The Digital Word: Text-based Computing in the Humanities*, 1993. Cambridge, MA: MIT Press.

[Lavagnino (1996)] John Lavagnino. 'Completeness and Adequacy in Text Encoding'. Richard J. Finneran (ed.) *The Literary Text in the Digital Age*, 1996. Ann Arbor, MI: University of Michigan Press. p. 63–76.

[Lightfoot (1979)] Charles Lightfoot. *Generic Textual Element Identification—A Primer*, 1979. Arlington: Graphic Communications Computer Association.

[Lubell (1999)] Joshua Lubell. 'Structured Markup on the Web: A Tale of Two Sites'. http://www.mel.nist.gov/msidlibrary/doc/mlang/markuplang.htm *Markup Languages: Theory & Practice* 1999. 1 (3) p. 7–22.

[McEnery et al. (1998)] Tony McEnery, Lou Burnard, Andrew Wilson, Paul Baker. *Validation of Linguistic Corpora*, http://users.ox.ac.uk/~lou/wip/ELRA/WP3/ 1998. (Report commissioned by ELRA)

[McGann (1997)] Jerome McGann. 'The Rationale of Hypertext'. Kathryn Sutherland (ed.) *Electronic Text: Investigations in Method and Theory*, 1997. New York, NY: Clarendon Press Oxford. p. 19–46.

[McGann (2001)] Jerome McGann. *Radiant Textuality: Literature After the World Wide Web*, 2001. New York, NY: Palgrave Macmillian.

[McGann (2004)] Jerome McGann. 'Marking Texts of Many Dimensions'. http://www.digitalhumanities.org/companion/ Susan Schreibman, Ray Siemens, John Unsworth (eds.) *A Companion to Digital Humanities*, 2004. Oxford: Blackwell. p. 198–217.

[Morrison et al. ((no date))] Alan Morrison, Michael Popham, Karen Wikander. *Creating and Documenting Electronic Texts: A Guide to Good Practice*, http://ota.ox.ac.uk/documents/creating/cdet/ (no date).

[Moser et al. (eds.) (2001)] Stephan Moser, Peter Stahl, Werner Wegstein, Norbert Richard Wolf (eds.) *Maschinelle Verarbeitung Altdeutscher Texte V (Beiträge zum Fünften Internationalen Symposion, Würzburg, 4–6 März 1997)*, 2001. Tübingen: Niemeyer.

[Pichler (1995)] Alois Pichler. 'Advantages of a Machine-Readable Version of Wittgenstein's Nachlaß'. http://hdl.handle.net/1956/1875 Kjell S. Johannessen, Tore Nordenstam (eds.) *Culture and Value: Philosophy and the Cultural Sciences. Beiträge des 18. Internationalen Wittgenstein Symposiums 13–20. August 1995 Kirchberg am Wechsel*, 1995. Kirchberg am Wechsel: Die Österreichische Ludwig Wittgenstein Gesellschaft. p. 770–776.

[Piez (2001)] Wendell Piez. 'Beyond the 'Descriptive vs. Procedural' Distinction'. http://conferences.idealliance.org/extreme/html/2001/Piez01/EML2001Piez01.html B. Tommie Usdin, Steven R. Newcomb (eds.) *Proceedings of Extreme Markup Languages 2001: Montreal, Canada*, 2001.

[Popham (1996)] Michael Popham. 'What Is Markup and Why Does It Matter'. Michael Popham, Lorna Hughes (eds.) *Computers and Teaching in the Humanities: Selected Papers from the CATH94 Conference held in Glasgow University September 9th-12th 1994*, 1996. Oxford: CTI Centre for Textual Studies.

[Quin (1996)] Liam Quin. 'Suggestive Markup: Explicit Relationships in Descriptive and Prescriptive DTDs'. https://www.holoweb.net/liam/papers/1996-sgml96-SuggestiveMarkup/ B. Tommie Usdin, Deborah A. Lapeyre (eds.) *SGML'96 Conference Proceedings*, 1996. Alexandria, VA: Graphic Communications Association. p. 405–418.

[Raymond et al. (1996)] Darrell Raymond, Frank Tompa, Derick Wood. 'From Data Representation to Data Model: Meta-Semantic Issues in the Evolution of SGML'. https://cs.uwaterloo.ca/~fwtompa/.papers/sgml.ps, http://www.sciencedirect.com/science/article/pii/0920548996000335, and http://dl.acm.org/citation.cfm?id=1648954 *Computer Standards & Interfaces* 1996. 18 (1) p. 25–36.

[Renear et al. (1996)] Allen Renear, David Durand, Elli Mylonas. 'Refining our Notion of What Text Really Is: The Problem of Overlapping Hierarchies'. http://cds.library.brown.edu/resources/stg/monographs/ohco.html Susan Hockey, Nancy Ide (eds.) *Research in Humanities Computing 4: Selected Papers from the 1992 ALLC/ACH Conference*, 1996. Oxford: Oxford University Press. p. 263–280.

[Renear (1997)] Allen Renear. 'Out of Praxis: Three (Meta)Theories of Textuality'. Kathryn Sutherland (ed.) *Electronic Text: Investigations in Method and Theory*, 1997. New York, NY: Clarendon Press Oxford. p. 107–126.

[Renear (2000)] Allen Renear. 'The Descriptive/Procedural Distinction is Flawed'. *Markup Languages: Theory and Practice* 2000. 2 (4) p. 411–420.

[Renear et al. (2002)] Allen H. Renear, David Dubin, C. Michael Sperberg-McQueen. 'Towards a Semantics for XML Markup'. 10.1145/585058.585081. Richard Furuta, Jonathan I. Maletic, Ethan V. Munson (eds.) *Proceedings of the 2002 ACM Symposium on Document Engineering*, 2002. McLean, VA: Association for Computing Machinery. p. 119–126.

[Renear et al. (2003)] Allen H. Renear, Christopher Phillippe, Pat Lawton, David Dubin. 'An XML Document Corresponds to Which FRBR Group 1 Entity?'. http://conferences.idealliance.org/extreme/html/2003/Lawton01/EML2003Lawton01.html B. Tommie Usdin, Steven R. Newcomb (eds.) *Proceedings of Extreme Markup Languages 2003: Montreal, Canada*, 2003.

[Renear et al. (2003)] Allen H. Renear, David Dubin, C. Michael Sperberg-McQueen, Claus Huitfeldt. 'XML Semantics and Digital Libraries'. http://dl.acm.org/citation.cfm?id=827192 *Proceedings of the 3rd ACM/IEEE–CS Joint Conference on Digital Libraries*, 2003. Los Alamitos, CA: IEEE Computer Society. p. 303–305.

[Renear (2004)] Allen H. Renear. 'Text Encoding'. http://www.digitalhumanities.org/companion/ Susan Schreibman, Ray Siemans, John Unsworth (eds.) *A Companion to Digital Humanities*, 2004. Oxford: Blackwell. p. 218–239.

[Salmon-Alt (2006)] Susanne Salmon-Alt. 'Data Structures for Etymology: Towards an Etymological Lexical Network'. http://hal.archives-ouvertes.fr/docs/00/11/09/71/PDF/etymology_final_bulag.pdf *BULAG: revue internationale annuelle Numéro Etymologie* 2006. Besançon: Presses Universitaires de Franche-Comté. 31.

[Schreibman (2002)] Susan Schreibman. 'Computer-mediated Texts and Textuality: Theory and Practice'. 10.1023/A:1016178200469. *Computers and the Humanities* 2002. 36 (3) p. 283–293.

[Schreibman (2002)] Susan Schreibman. 'The Text Ported'. 10.1093/llc/17.1.77. *Literary and Linguistic Computing* 2002. 17 (1) p. 77–87.

[SGML Users' Group (1990)] SGML Users' Group. *A Brief History of the Development of SGML*, http://www.sgmlsource.com/history/sgmlhist.htm 1990.

[Shipman and Marshall (1999)] Frank M. ShipmanIII, Catherine C. Marshall. 'Formality Considered Harmful: Experiences, Emerging Themes, and Directions on the Use of Formal Representations in Interactive Systems'. 10.1023/A:1008716330212. http://www.csdl.tamu.edu/~shipman/papers/cscw.pdf *Computer-Supported Cooperative Work* 1999. 8 (4) p. 333–352.

[Sperberg-McQueen and Huitfeld (1999)] C. Michael Sperberg-McQueen, Claus Huitfeld. 'Concurrent Document Hierarchies in MECS and SGML'. *Literary and Linguistic Computing* 1999. 14 (1) pp. 29-42.

[Sperberg-McQueen and Huitfeldt (1999)] C. Michael Sperberg-McQueen, Claus Huitfeldt. 'Concurrent document hierarchies in MECS and SGML'. 10.1093/llc/14.1.29. *Literary and Linguistic Computing* 1999. 14 (1) p. 29–42.

[Sperberg-McQueen et al. (2000)] C. Michael Sperberg-McQueen, Claus Huitfeldt, Allen H. Renear. 'Meaning and Interpretation in Markup'. *Markup Languages: Theory and Practice* 2000. 2 (3) p. 215–234.

[Sperberg-McQueen et al. (2002)] C. Michael Sperberg-McQueen, David Dubin, Claus Huitfeldt, Allen Renear. 'Drawing Inferences on the Basis of Markup'. http://conferences.idealliance.org/extreme/html/2002/CMSMcQ01/EML2002CMSMcQ01.html B. Tommie Usdin, Steven R. Newcomb (eds.) *Proceedings of Extreme Markup Languages 2002: Montreal, Canada*, 2002.

[Sperberg-McQueen (2006)] C. Michael Sperberg-McQueen. 'Rabbit/duck grammars: a validation method for overlapping structures'. http://conferences.idealliance.org/extreme/html/2006/SperbergMcQueen01/EML2006SperbergMcQueen01.html *Proceedings of Extreme Markup Languages 2006*, 2006.

[Sukovic (2002)] Suzana Sukovic. 'Beyond the Scriptorium: The Role of the Library in Text Encoding'. http://www.dlib.org/dlib/january02/sukovic/01sukovic.html *D-Lib* 2002. 8 (1) .

[University of Nebraska — Lincoln Libraries (2003)] University of Nebraska — Lincoln Libraries. *A Basic Guide to Text Encoding*, http://cdrh.unl.edu/book/export/html/2428 2003.

[Unsworth (2000)] John Unsworth. *Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, How Might Our Tools Reflect This?*, http://people.brandeis.edu/~unsworth/Kings.5-00/primitives.html 2000. (Part of a Symposium on "Humanities Computing: Formal Methods, Experimental Practice" sponsored by King's College, London)

[Unsworth (2001)] John Unsworth. *Knowledge Representation in Humanities Computing*, http://people.brandeis.edu/~unsworth/KR/ 2001. (Lecture I in the eHumanities NEH Lecture Series on Technology & the Humanities, Washington, DC, April 3, 2001)

[Vitali et al. (2000)] Fabio Vitali, Luca Bompani, Paolo Ciancarini. 'Hypertext Functionalities with XML'. *Markup Languages: Theory & Practice* 2000. 2 (4) p. 389.

[Watson (1992)] Dennis G. Watson. *Brief History of Document Markup*, http://chnm.gmu.edu/digitalhistory/links/pdf/chapter3/3.19a.pdf 1992. (Circular 1086. Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida)

[Weel ((no date))] Adriaan Weel. 'The Concept of Markup'. https://docmh.com/adriaan-van-der-weel-digital-text-and-the-gutenberg-heritage-pdf *Digital Text and the Gutenberg Heritage*, (no date). chapter 3. (in preparation; draft only)

[Welty and Ide (1999)] Christopher Welty, Nancy Ide. 'Using the Right Tools: Enhancing Retrieval from Marked-up Documents'. 10.1023/A:1001800717376. http://link.springer.com/article/10.1023/A%3A1001800717376 *Computers and the Humanities* 1999. 33 (1–2) p. 59–84.

## TEI

[An Agreement to Establish a Consortium for the Maintenance of the Text Encoding Initiative (March 1999)] *An Agreement to Establish a Consortium for the Maintenance of the Text Encoding Initiative*, http://www.tei-c.org/About/consortium.html March 1999.

[Bauman (1995)] Syd Bauman. 'Tables of Contents TEI-style'. Lou Burnard (ed.) *TEXT Technology: The Journal of Computer Text Processing Electronic Texts and the Text Encoding Initiative. A Special Issue of TEXT Technology* 1995. Madison, SD: College of Liberal Arts, Dakota State University. 5 (3) p. 235–247.

[Bauman (1996)] Syd Bauman. 'Keying NAMEs: the WWP Approach'. http://www.wwp.brown.edu/about/history/archive/newsletter/vol02num03/nameKey-home.html *Brown University Women Writers Project Newsletter* 1996. 2 (3) p. 3–6 p. 10–11.

[Bauman and Catapano (1999)] Syd Bauman, Terry Catapano. 'TEI and the Encoding of the Physical Structure of Books'. 10.1023/A:1001769103586. *Computers and the Humanities* 1999. 33 (1–2) p. 113–127.

[Bauman and Flanders (2004)] Syd Bauman, Julia Flanders. 'Odd Customizations'. http://conferences.idealliance.org/extreme/html/2004/Bauman01/EML2004Bauman01.html *Proceedings of Extreme Markup Languages 2004*, 2004.

[Bauman (2005)] Syd Bauman. 'TEI HORSEing Around'. http://conferences.idealliance.org/extreme/html/2005/Bauman01/EML2005Bauman01.html *Proceedings of the Extreme Markup Languages 2005*, 2005.

[Brown (1994)] Malcolm B. Brown. 'What is the TEI?'. *Information Technology and Libraries* 1994. 13 (1) p. 8.

[Burnard (1992)] Lou Burnard. 'The Text Encoding Initiative: A Progress Report'. Gerhard Leitner (ed.) *New Directions in Corpus Linguistics*, 1992. Berlin: Mouton de Gruyter.

[Burnard (1993)] Lou Burnard. 'Rolling your own with the TEI'. *Information Services and Use* 1993. Amsterdam: IOS Press. 13 (2) p. 141–154.

[Burnard (1994)] Lou Burnard. 'The TEI: Towards an Extensible Standard for the Encoding of Texts'. Seamus Ross, Edward Higgs (eds.) *Electronic Information Resources and Historians*, 1994. London: British Academy.

[Burnard (1995)] Lou Burnard. 'The Text Encoding Initiative: An Overview'. Geoffrey Leech, Greg Myers, Jenny Thomas (eds.) *Spoken English on Computer: Transcription, Mark-up and Application*, 1995. London: Longman.

[Burnard (1997)] Lou Burnard. *The Text Encoding Initiative's Recommendations for the Encoding of Language Corpora: Theory and Practice*, http://users.ox.ac.uk/~lou/wip/Soria/ 1997. (Prepared for a seminar on Etiquetación y extracción de información de grandes corpus textuales within the Curso Industrias de la Lengua (14–18 de Julio de 1997). Sponsored by the Fundacion Duques de Soria.)

[Burnard and Popham (1999)] Lou Burnard, Michael Popham. 'Putting Our Headers Together: A Report on the TEI Header Meeting 12 September 1997'. 10.1023/A:1001710828622. *Computers and the Humanities* 1999. Dordrecht, Boston: Kluwer Academic Publishers. 33 (1-2) p. 39–47.

[Burnard (2000)] Lou Burnard. *Text Encoding for Interchange: A New Consortium*, http://www.ariadne.ac.uk/issue24/tei 2000.

[Chang (2001)] Sheau-Hwang Chang. 'The Implications of TEI'. *OCLC Systems and Services* 2001. 17 (3) p. 101–103.

[Ciotti (ed.) (2005)] Fabio Ciotti (ed.) *Il Manuale TEI Lite: Introduzione Alla Codifica Elettronica Dei Testi Letterari*, 2005. Milano: Sylvestre Bonnard.

[Cournane (1997)] Mavis Cournane. *The Application of SGML/TEI to the Processing of Complex, Multi-lingual Text*, (PhD Dissertation)1997. Cork, Ireland: University College Cork.

[Digital Library Federation (1998)] Digital Library Federation. *TEI and XML in Digital Libraries: Meeting June 30 and July 1, 1998, Library of Congress, Summary/Proceedings*, 1998.

[Digital Library Federation (2011)] Digital Library Federation. *TEI Text Encoding in Libraries: Guidelines for Best Encoding Practices, Version 3.0 (October 2011)*, http://www.tei-c.org/SIG/Libraries/teiinlibraries/main-driver.html 2011.

[Finney (2006)] Timothy J. Finney. 'Manuscript Markup'. Larry W. Hurtado (ed.) *The Freer Biblical Manuscripts: Fresh Studies of an American Treasure Trove*, 2006. Atlanta, GA: Society of Biblical Literature. pp. 263-288. Text-critical studies 6.

[Gibson and Ruotolo (2003)] Matthew Gibson, Christine Ruotolo. 'Beyond the Web: TEI, the Digital Library, and the Ebook Revolution'. 10.1023/A:1021895322291. *Computers and the Humanities* 2003. 37 (1) p. 57–63.

[Kindberg and Hawke (eds.) (2005)] T. Kindberg, S. Hawke (eds.) *The 'tag' URI Scheme*, https://www.ietf.org/rfc/rfc4151.txt RFC 4151. 2005. IETF.

[Loiseau (2002)] Sylvain Loiseau. *Les standards : autour d'XML et de la TEI*, http://www.revue-texto.net/Corpus/Manufacture/standards/sommaire.html 2002.

[Marko and Kelleher Powell (2001)] Lynn Marko, Christina Kelleher Powell. 'Descriptive Metadata Strategy for TEI Headers: A University of Michigan Library Case Study'. 10.1108/10650750110402585. *OCLC Systems & Services* 2001. 17 (3) pp. 117-20.

[Mertz (2003)] David Mertz. *XML Matters: TEI — the Text Encoding Initiative, An XML Dialect for Archival and Complex Documents*, http://www.ibm.com/developerworks/xml/library/x-matters30/index.html 2003.

[Morrison (1999)] Alan Morrison. 'Delivering Electronic Texts Over the Web: The Current and Planned Practices of the Oxford Text Archive'. 10.1023/A:1001726011322. *Computers and the Humanities* 1999. 33 (1-2) pp. 193-198.

[Mylonas and Renear (1999)] Elli Mylonas, Allen Renear. 'The Text Encoding Initiative at 10: Not Just an Interchange Format Anymore — But a New Research Community'. 10.1023/A:1001832310939. *Computers and the Humanities* 1999. 33 (1-2) pp. 1-9.

[Nellhaus (2001)] Tobin Nellhaus. 'XML, TEI, Digital Libraries in the Humanities'. http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v001/1.3nellhaus.html *Portal: Libraries and the Academy* 2001. 1 (3) pp. 267-277.

[Rahtz (2003)] Sebastian Rahtz. *Building TEI DTDs and Schemas on demand*, https://web.archive.org/web/20150706100318/http://tei.oucs.ox.ac.uk/Talks/2003-05-06-xmleurope2003/xmleurope2003.pdf 2003. (Paper presented at XML Europe 2003, London, March 2003)

[Rahtz et al. (2004)] Sebastian Rahtz, Norman Walsh, Lou Burnard. *A unified model for text markup: TEI, Docbook, and beyond*, http://www.tei-c.org/Activities/Workgroups/META/xmleurope2004.pdf 2004. (Paper presented at XML Europe 2004, Amsterdam, April 2004)

[Renear (1995)] Allen Renear. 'Theory and Metatheory in the Development of Text Encoding'. http://uhra.herts.ac.uk/bitstream/handle/2299/7302/904841.pdf?sequence=2 Michael A. R. Biggs, Claus Huitfeldt (eds.) *Philosophy and Electronic Publishing*, 1995. (Interactive seminar for the Monist)

[Robinson ((no date))] Peter Robinson. *Making a Digital Edition with TEI and Anastasia*, http://sd-editions.com/AnaServer?teidoc+0+start.anv (no date).

[Seaman (1995)] David Seaman. *The Electronic Text Center Introduction to TEI and Guide to Document Preparation*, https://web.archive.org/web/20140804000849/etext.lib.virginia.edu/standards/tei/uvatei.html 1995.

[Simons (1999)] Gary F. Simons. 'Using Architectural Forms to Map TEI Data into an Object-Oriented Database'. 10.1023/A:1001765030032. *Computers and the Humanities* 1999. 33 (1-2) pp. 85-101.

[Smith (1999)] David Smith. 'Textual Variation and Version Control in the TEI'. 10.1023/A:1001795210724. *Computers and the Humanities* 1999. 33 (1-2) pp. 103-112.

[Sperberg-McQueen (1991)] C. Michael Sperberg-McQueen. 'Text in the Electronic Age: Textual Study and Text Encoding, with Examples from Medieval Texts'. 10.1093/llc/6.1.34. *Literary & Linguistic Computing* 1991. 6 (1) pp. 34-46.

[Sperberg-McQueen (1994)] C. Michael Sperberg-McQueen. 'The Text Encoding Initiative: Electronic Text Markup for Research'. Brett Sutton (ed.) *Literary Texts in an Electronic Age*, 1994. Urbana-Champaign, IL: University of Illinois at Urbana-Champaign, Graduate School of Library and Information Science. p. 35–55.

[Sperberg-McQueen (1996)] C. Michael Sperberg-McQueen. 'Textual Criticism and the Text Encoding Initiative'. Richard J. Finneran (ed.) *The Literary Text in the Digital Age*, 1996. Ann Arbor, MI: University of Michigan Press. p. 37–62.

[Vanhoutte (2004)] Edward Vanhoutte. 'An Introduction to the TEI and the TEI Consortium'. 10.1093/llc/19.1.9. *Literary & Linguistic Computing* 2004. 19 (1) p. 9.

# G Deprecations

## G.1 Items Scheduled to be Removed

This appendix provides a partial list of items (elements, attributes, attribute values, datatypes etc.) which are deprecated in this revision of the TEI Guidelines (revision ebf0834e1 of TEI P5 Version 4.10.0).

- Because an <entry> can now occur inside an <entry>, the <re> element is no longer needed. These Guidelines recommend using an <entry> with a *type* attribute value of relatedEntry instead. will be removed on 2026-03-10.
- Because an <entry> can now occur inside an <entry>, the <superEntry> element is no longer needed. For instances where the <superEntry> element was used to group multiple <entry> elements, these Guidelines recommend using instead an outer <entry> element, perhaps with a *type* attribute. will be removed on 2027-03-07.

# H Prefatory Notes

This Appendix contains (in reverse chronological order) the 'Introductory Notes' prefixed to each revision of the TEI Guidelines since its first publication in 1994.

## H.1 Prefatory Note (November 2007)

This publication constitutes the fifth distinct version of the *Guidelines for Electronic Text Encoding and Interchange*, and the first complete revision since the appearance of P3 in 1994. It includes substantial amounts of new material and a major revision of the underlying technical infrastructure. With this version, these Guidelines enter a new stage in their development as a community-maintained open source project. This edition is the first version to have benefitted from the close overview and oversight of an elected TEI Technical Council. The editors are therefore particularly pleased to acknowledge with gratitude the hard work and dedication put into this project by the Council over the last five years.

The Chair of the TEI Board sits on the Technical Council, and the Board appoints the Chair of the Technical Council and one other member of the Council. Other Council members are all elected by the Consortium membership, and serve periods of up to two years at a time. *[Note: What followed was a list of TEI Board and Council members, which was periodically updated over the years, and remains in place with the 2025 Introductory Note.]*

## H.2 Prefatory Note (March 2002)

The primary goal of this revision has been to make available a new and corrected version of the TEI Guidelines which:

- is expressed in XML and conforms to a TEI-conformant XML DTD;
- generates a set of DTD fragments that can be combined together to form either SGML or XML document type definitions;
- corrects blatant errors, typographical mishaps, and other egregious editorial oversights;
- can be processed and maintained using readily available XML tools instead of the special-purpose ad hoc software originally used for TEI P3.

A second major design goal of this revision has been to ensure that the DTD fragments generated would not break existing documents: in other words, that any document conforming to the original TEI P3 SGML DTD would also conform to the new XML version of it. Although full backwards compatibility cannot be guaranteed, we believe our implementation is consistent with that goal.

In most respects, the TEI Guidelines have stood the test of time remarkably well. The present edition makes no substantial attempt to rewrite those few parts of them which have now been rendered obsolete by changes since their first publication, though an indication is given in the text of where such rewriting is now considered necessary. Neither does the present version attempt to address any of the many possible new areas of digital activity in which the TEI approach to standardization may have something to offer. Both these tasks require the existence of an informed and active TEI Council to direct and validate such extension and maintenance work, in response to the changing needs and priorities of the TEI user community.

Two exceptions to the above principles may be cited: firstly, the chapter which originally provided a 'Gentle Introduction' to SGML has been completely rewritten to provide a similarly gentle introduction to XML; secondly the chapter on character sets has been completely revised in light of the close connexion between Unicode and XML. The editors gratefully acknowledge the assistance of the *ad hoc* workgroup chaired by Christian Wittern, which undertook to provide expert advice and correction at very short notice, in the latter task.

The preparation of this new version relied extensively on preliminary work carried out by the former North American editor of the TEI Guidelines, C.M. Sperberg-McQueen. In a TEI working paper written in 1999[143] he sketched out a precise blueprint for the conversion of the TEI from SGML to XML, which we have implemented, with only slight modification.

The Editors would also like to express thanks to the team of volunteers from the TEI community who helped us with the task of proofreading the first draft during the summer of 2001; and to Sebastian Rahtz of Oxford University Computing Services, without whose skill and enthusiasm this new edition would not have been possible.

[143] *TEI ED W69*, available from the TEI web site at `http://www.tei-c.org/Vault/ED/edw69.htm`.

A substantial proportion of the work of preparing this new edition was funded with the assistance of a grant from the US National Endowment for the Humanities, whose continued support of the TEI has also been crucial to the effort of setting up the TEI Consortium.

Finally, we would like to thank all our colleagues on the interim management board of the TEI Consortium, in particular its Chairman John Unsworth, for their continued support of the TEI's work, and their willingness to devote effort to the difficult task of overseeing its transition to a new organizational infrastructure.

Summary details of the changes made in the present and previous editions are given in their Prefatory Notes, all of which are now reproduced in an Appendix to the present edition: see *Appendix H Prefatory Notes*.

Lou Burnard and Syd Bauman (TEI Editors)
Oxford and Providence, March 2002.

## H.3 Introductory Note (November 2001)

To complete the work started in June of this year, the TEI Editors asked for volunteers from the TEI community to proofread the preliminary XML version. 24 volunteers responded to this call during August, and gave invaluable help both by identifying a number of previously un-noticed errors, and by suggesting areas in which more substantial revision should be undertaken in the future. The Editors gratefully acknowledge the assistance of the following individuals during this exercise:

Jimmy Adair, Syd Bauman, Michael Beddow, Steven Bird, Lisa Charlong, Matthew Driscoll, Patrick Durusau, Tomaz Erjavec, Nick Finke, Tim Finney, Julia Flanders, Mike Fraser, Pankaj Kamthan, François Lachance, Terry Langendoen, Anne Mahoney, Gregory Murphy, Daniel Pitti, Rafal Prinke, Laurent Romary, Stewart Russell, Gary Simons, Elisabeth Solopova, Christian Wittern, Martin Wynne.

In addition to error correction, and clear delineation of those sections in which substantial revision is yet to be undertaken for TEI P5, the present draft differs from earlier ones in the following respects:

- Formal Public Identifiers have been introduced as a means of constructing TEI DTDs and an SGML Open Catalog is now included with the standard release;
- Some systematic errors and omissions in the reference section have been removed; the format of this section has been substantially changed, we hope for the better;
- The chapters on obtaining the TEI DTDs and WSDs have been brought up to date; the chapter on modification has been expanded to include a discussion of the TEI Lite customization;
- All examples and cited markup has been checked for XML validity against the published DTDs, and corrected where faulty; examples have been formatted in a (more or less) consistent style.

Lou Burnard and Syd Bauman (Editors)
Oxford and Providence, November 2001.

## H.4 Introductory Note (June 2001)

This is a preliminary version of a revised and fully XML-compliant edition of the TEI Guidelines. Although work on revising and correcting the text of the document is incomplete, by making available this preliminary version we hope to facilitate testing of the XML document type declarations which it describes by as wide a range of TEI users as possible.

The primary goal of this revision is to make available the corrected (May 1999) edition of the Guidelines in a new version which:

- is expressed in XML and itself conforms to a TEI-conformant XML DTD;
- generates a set of XML DTD fragments that can be combined together in the same way as the existing TEI (P3) SGML DTD fragments to form true TEI XML DTD fragments without loss of functionality;
- can be processed and maintained using readily available XML tools instead of the special-purpose ad hoc software originally used for TEI P3.

As noted elsewhere, a number of errors were corrected in the May 1999 edition. A (much) smaller number of errors have also been corrected in this edition, but no new material has been added. We expect the expansion

and modification of the Guidelines to become a real possibility in the context of the newly formed TEI Consortium, which has funded the preparation of this present edition.

A major design goal of both this and the previous revision has been to ensure that the DTD fragments generated would not break existing documents: in other words, that any document conforming to the original TEI P3 SGML DTD would also conform to the new XML version of it. Although full backwards compatibility cannot be guaranteed, we believe our implementation is consistent with that goal.

In making this new version, we relied extensively on preliminary work carried out by the outgoing North American editor of the TEI Guidelines, Michael Sperberg-McQueen. In a TEI working paper written in 1999, TEI ED W69, Michael sketched out a precise blueprint for the conversion of the TEI from SGML to XML, which we have implemented, with only slight modification. The current TEI editors wish to express here our admiration for the detailed care put into that paper, without which our task would have been forbiddingly difficult, if not impossible. We would also like to express our thanks to Sebastian Rahtz of Oxford University Computing Services, for his invaluable assistance in preparing this new edition.

We list here in summary form all the changes made in the present edition. Full technical details are provided in documents TEI EDW69 and TEI EDW70, available from the TEI web site.

1. A new keyword `TEI.XML` has been added. By setting its value to INCLUDE, rather than the default IGNORE, the user can request generation of an XML rather than an SGML DTD;
2. The content models of all elements have been checked, and, where necessary, changed so that they are equally valid as SGML or as XML;
3. The declared value for all attributes has been changed to a form which is equally valid as SGML or as XML;
4. All the examples have been checked for conformance and converted to use XML syntax, where possible. (This process is currently incomplete.)
5. Some errors and duplications in the class membership of elements from the names and dates tagsets have been corrected.

To implement the first of these, we have parameterized the *tag omissibility* indicators '- o' and '- -' used within element declarations in the DTD. When XML is to be generated, the parameter entities concerned are redeclared with the null string as their value.

The second change was achieved by removing SGML-specific features (ampersand connectors, inclusion and exclusion exceptions, various types of attribute content) from the DTD and revising the syntax of the DTD to conform to XML requirements (specifically in the representation of mixed-content models, and by removing redundant parentheses). In making these changes, we took care to ensure that the resulting content model would continue to accept existing valid documents, though in the nature of things it could not be guaranteed to reject the same set of documents. As further discussed in EDW69 and EDW70, some constraints (exclusion exceptions, for example) which could be carried out by a generic SGML parser using TEI P3 will have to be implemented by a special purpose TEI validator using TEI P4.

Much work remains to be done, firstly in testing the new DTD fragments against as wide a range of TEI materials as possible, secondly in revising the discussion of markup theory and practice within the text to reflect current thinking. A few sections of the current text (the Gentle Introduction to SGML and the discussion of Extended Pointer syntax are two examples) will need substantial rewriting. For the most part, however, we think the Guidelines have stood the test of time well and can be recommended to a new generation of text encoders scarcely born at the time they were first formulated.

Lou Burnard and Steve De Rose (Editors)
Oxford and Providence, May 2001.

## H.5 Introductory Note (May 1999)

No work of the size and complexity of the TEI *Guidelines* could reasonably be expected to be error-free on publication, nor to remain long uncorrected. It has however taken rather longer than might have been anticipated to complete production of the present corrected reprint of the first edition, for which we present our apologies, both to the many individuals and institutions whose enthusiastic adoption and promotion of the TEI encoding scheme have ensured its continued survival in the rapidly changing world of digital scholarship, and also to the many helpfully critical users whose assiduous uncovering and reporting of our errors have made possible the present revision.

At its first meeting in Bergen, in June 1996, the TEI Technical Review Committee (TRC) approved the setting up of a small working committee to oversee the production of a revised edition of the TEI *Guidelines*, to include corrections of as many as possible of the 'corrigible errors' notified to the editors since publication of the first edition in May 1994, the bulk of which are summarized in a TEI working paper (TEI EDW67, available from the TEI web site).

During the spring of 1997, this TRC Core Subcommittee reviewed nearly 200 comments and proposals which the editors had collected from public debate and discussion over the preceding two years, and provided invaluable technical guidance in disposition of them. We are glad to take this opportunity of expressing our thanks to this subcommittee, whose members were Elli Mylonas, Dominic Dunlop, and David T. Barnard.

The work of making the corrections and regenerating the text proceeded rather fitfully during 1998 and 1999, largely because of increasing demands on the editors' time from their other responsibilities. With the establishment of the new TEI Consortium, it is be hoped that maintenance of the Guidelines will be placed on a more secure footing. Some specific areas in which we anticipate future revisions being carried out are listed below.

## Typographic Corrections Made

- examples of TEI markup throughout the text were all checked against the relevant DTD fragment and an embarassingly large number of tagging errors corrected;
- various minor typographic and spelling errors were corrected;
- the 'corrigible errors' listed in working paper TEI EDW67 were all corrected: some of these required specific changes to the DTD which are listed in the next section.

## Specific Changes in the DTD

A major goal of this revision was to avoid changes which might invalidate existing data, even where existing constructs seemed erroneous in retrospect. To that end, wherever changes have been made in content models for existing elements, they have as far as possible been made so that the DTD will now accept a superset of what was previously legal. Only one new element (<ab>) has been added.

Where possible, a few content models have been changed in such a way as to facilitate conversion to XML, but XML compatibility is *not* a goal of this revision.

Brief details of all changes made in the DTD follow:

- Several changes were made in class membership, in order to correct unreachability problems. Specifically:
  - elements <geogName>, <persName>, <placeName> were added to the `m.data` class;
  - <geogName> and <placeName> were removed from the `m.placepart` class;
  - the elements <addSpan>, <delSpan>, <gap>, were added to the `m.Edit` class;
  - a new class `m.editIncl` was defined, with members <addSpan>, <delSpan>, and <gap>; this class was then added to the global inclusion class `m.globIncl` along with <anchor> (erroneously a member of the `m.Seg` class, from which it is now removed), `m.metadata` and `m.refsys`;
- added <name> element to `m.addrPart` class;
- added <dateline> to `m.divtop` and `m.divbot` classes;
- added <epilogue> and <castList> to `m.dramafront` class;
- added <divGen> to `m.front` class;
- added <dateline> to `m.divtop` and `m.divtop` classes;
- added <u> element to `a.declaring` class;
- defined new class `m.fmchunk` (front matter chunk), comprising <argument>, <byline>, <docAuthor>, <docDate>, <docEdition>, <docImprint>, <docTitle>, <epigraph>, <head>, and <titlePart> for use in simplification of the content model for <front> element;
- defined new element <ab> (anonymous block), and added it to the `m.chunk` class;
- corrected an error whereby global attributes were not properly defined for elements specifying a non-default value for any of the `a.global` attributes: elements affected include: <foreign>, <hi>, <del>, <pb>, <lb>, <cb>, <language>, <anchor>, and <when>;
- changed content models to permit empty <list> and empty <availability> elements;

- changed content model for <series> element to permit #PCDATA;
- changed content model for <setting> element to permit <date> element as a direct child;
- added a *key* attribute to the `<distance>` element, for consistency with other elements in its class;
- changed content model for <orgName> element to make it more consistent with e.g. <persName>;
- changed content model for <opener> element to include <argument>, <byline>, and <epigraph>;
- changed content models for <app>, <rdgGrp>, and <wit> elements;
- revised attributes on `<hand>` element.

A number of content models were changed with a view to easing the creation of an XML compatible version of the Guidelines. Specifically:

- removed ampersand connectors from <cit>, <respStmt>, <publicationStmt>, and <graph>;
- changed the mixed content models for <sense>, <re>, <persName>, <placeName>, <geogName>, `<dateStruct>`, `<timeStruct>`, and <dateline> to make them XML-conformant.

### Outstanding Errors

A small number of other known problems remain uncorrected in this version and are briefly listed below. Please watch the TEI mailing list for announcements of their correction.

- elements of class `model.inter` don't always behave as they should (e.g. one cannot insert a <table> before anything else in a );
- some mixed-content problems consequent on the definition of `macro.specialPara` need to be addressed systematically; in particular, the treatment of list items or notes which contain several paragraphs continues to surprise many users: no whitespace is allowed between the paragraphs;
- the *resp* attributes on editorial elements are not consistently defined;
- the discussions of DTD invocation, and the DTD itself, all use system identifiers instead of formal public identifiers.

Our next priority however will be the production of a fully XML-compliant version of the TEI DTD, work on which is already well advanced.

C.M. Sperberg-McQueen and Lou Burnard, May 1999

## H.6 Preface (April 1994)

These Guidelines are the result of over five years' effort by members of the research and academic community within the framework of an international cooperative project called the Text Encoding Initiative (TEI), established in 1987 under the joint sponsorship of the Association for Computers and the Humanities, the Association for Computational Linguistics, and the Association for Literary and Linguistic Computing.

The impetus for the project came from the humanities computing community, which sought a common encoding scheme for complex textual structures in order to reduce the diversity of existing encoding practices, simplify processing by machine, and encourage the sharing of electronic texts. It soon became apparent that a sufficiently flexible scheme could provide solutions for text encoding problems generally. The scope of the TEI was therefore broadened to meet the varied encoding requirements of any discipline or application. Thus, the TEI became the only systematized attempt to develop a fully general text encoding model and set of encoding conventions based upon it, suitable for processing and analysis of any type of text, in any language, and intended to serve the increasing range of existing (and potential) applications and use.

What is published here is a major milestone in this effort. It provides a single, coherent framework for all kinds of text encoding which is hardware-, software- and application-independent. Within this framework, it specifies encoding conventions for a number of key text types and features. The ongoing work of the TEI is to extend the scheme presented here to cover additional text types and features, as well as to continue to refine its encoding recommendations on the basis of extensive experience with their actual application and use.

We therefore offer these Guidelines to the user community for use in the same spirit of active collaboration and cooperation with which they have so far been developed. The TEI is committed to actively supporting the widespread and large-scale use of the Guidelines which, with the publication of this volume, is now for the first time

possible. In addition, we anticipate that users of the TEI Guidelines will in some instances adapt and extend them as necessary to suit particular needs; we invite such users to engage in the further development of these Guidelines by working with us as they do so.

Like any standard which is actually used, these Guidelines do not represent a static finished work, but rather one which will evolve over time with the active involvement of its community of users. We invite and encourage the participation of the user community in this process, in order to ensure that the TEI Guidelines become and remain useful in all sorts of work with machine-readable texts.

This document was made possible in part by financial support from the U.S. National Endowment for the Humanities, an independent federal agency; Directorate General XIII of the Commission of the European Communities; the Andrew W. Mellon Foundation; and the Social Science and Humanities Research Council of Canada. Direct and indirect support has also been received from the University of Illinois at Chicago, the Oxford University Computing Services, the University of Arizona, the University of Oslo and Queen's University (Kingston, Ont.), Bellcore (Bell Communications Research), the Istituto di Linguistica Computazionale (C.N.R.) Pisa, the British Academy, and Ohio State University, as well as the employers and host institutions of the members of the TEI working committees and work groups listed in the acknowledgments.

The production of this document has been greatly facilitated by the willingness of many software vendors to provide us with evaluation versions of their products. Most parts of this text have been processed at some time by almost every currently available SGML-aware software system. In particular, we gratefully acknowledge the assistance of the following vendors:

- Berger-Levrault AIS s.a. (for Balise);
- E2S n.v. (for E2S Advanced SGML Editor);
- Electronic Book Technology (for DynaText);
- SEMA Group and Yard Software (for Mark-It and Write-It);
- Software Exoterica (for CheckMark and Xtran);
- SoftQuad, Inc., (for Author/Editor and RulesBuilder);
- Xerox Corporation (for Ventura Publisher).

Details of the software actually used to produce the current document are given in the colophon at the end of the work.

## H.7 Acknowledgments

Many people have given of their time, energy, expertise, and support in the creation of this document; it is unfortunately not possible to thank them all adequately. Below are listed those who have served as formal members of the TEI's Work Groups and Working Committees during its six-year history; others not so officially enfranchised also contributed much to the quality of the result.

The editors take this opportunity to acknowledge our debt to those who have patiently endured and corrected our misunderstandings of their work; we hope that they will feel the wait has not been in vain. For any errors and inconsistencies remaining, we must accept responsibility; any virtue in what is here presented, we gladly ascribe to the energies of the keen intellects listed below.

C. M. Sperberg-McQueen and Lou Burnard

### TEI Working Committees (1990-1993)

Not all members listed were able to serve throughout the development of the Guidelines.

**Committee on Text Documentation**: Chair: Dominik Wujastyk (Wellcome Institute for the History of Medicine)

Members 1990–1992: J. D. Byrum (Library of Congress); Marianne Gaunt (Rutgers University); Richard Giordano (Manchester University); Barbara Ann Kipfer (Independent Consultant); Hans Jørgen Marker (Danish Data Archive, Odense); Marcia Taylor (University of Essex);

**Committee on Text Representation** Chair: Stig Johansson (University of Oslo)

Members 1990–1992: Roberto Cencioni (Commission of the European Communities); David R. Chesnutt (University of South Carolina); Robin C. Cover (Dallas Theological Seminary); Steven J. DeRose (Electronic Book Technology Inc); David G. Durand (Boston University); Susan M. Hockey (Oxford University Computing

Service); Claus Huitfeldt (University of Bergen); Francisco Marcos-Marin (University Madrid); Elli Mylonas (Harvard University); Wilhelm Ott (University of Tübingen); Allen H. Renear (Brown University); Manfred Thaller (Max-Planck-Institut für Geschichte, Göttingen)

**Committee on Text Analysis and Interpretation** Chair: D. Terence Langendoen (University of Arizona)

Members 1990–1992: Robert Amsler (Bell Communications Research); Stephen Anderson (Johns Hopkins University); Branimir Boguraev (IBM T. J. Watson Research Center); Nicoletta Calzolari (University of Pisa); Robert Ingria (Bolt Beranek Newman Inc); Winfried Lenders (University of Bonn); Mitch Marcus (University of Pennsylvania); Nelleke Oostdijk (University of Nijmegen); William Poser (Stanford University); Beatrice Santorini (University of Pennsylvania); Gary Simons (Summer Institute of Linguistics); Antonio Zampolli, University of Pisa.

**Committee on Metalanguage and Syntax** Chair: David T. Barnard (Queen's University)

Members 1990–1994: David G. Durand (Boston University); Jean-Pierre Gaspart (Associated Consultants and Software Engineers sa/nv); Nancy M. Ide (Vassar College); Lynne A. Price (Software Exoterica / Xerox PARC); Frank Tompa (University of Waterloo); Giovanni Battista Varile (Commission of the European Communities).

In addition, the two TEI editors served ex officio on each committee.

Following publication of the first draft of the TEI Guidelines (P1) in November 1990, a number of specialist work groups were charged with responsibility for drafting revisions and extensions, which, together with material already presented in P1, constitute the basis of the present work.

In addition, many members of the work groups listed below met on three occasions to review the emerging proposals in detail at technical review meetings convened by the TEI Steering Committee. These meetings, held in Myrdal, Norway (November 1991), Chicago (May 1992) and Oxford (May 1993), were largely responsible for the technical content and organization of the present work. Attendants at these meetings are starred in the list below.

**TR1 Character sets** Chair: Harry Gaylord* (University of Groningen); Syun Tutiya* (Chiba University).

**TR2 Text criticism** Chair: Peter Robinson* (Oxford University); David Chesnutt* (University of South Carolina); Robin Cover* (Dallas Theological Seminary); Robert Kraft (University of Pennsylvania); Peter Shillingsburg (Mississippi State University).

**TR3 Hypertext and hypermedia** Chair: Steven J. DeRose* (Electronic Book Technologies Inc); David Durand (Boston University); Edward A. Fox (Virginia State University); Eve Wilson (University of Kent).

**TR4 Formulæ, Tables, figures, and graphics** Chair: Paul Ellison* (University of Exeter); Anders Berglund (Independent Consultant); Dale Waldt (Thompson Professional Publishing).

**TR6 Language corpora** Chair: Douglas Biber* (University of Northern Arizona); Jeremy Clear (Birmingham University); Gunnel Engwall (University of Stockholm).

**TR9 Manuscripts and codicology** Chair: Claus Huitfeldt* (University of Bergen); Dino Buzzetti (University of Bologna); Jacqueline Hamesse (University of Louvain); Mary Keeler (Georgetown University); Christian Kloesel (Indiana University); Allen Renear* (Brown University); Donald Spaeth (Glasgow University).

**TR10 Verse** Chair: David Robey* (University of Manchester); Elaine Brennan* (Brown University); David Chisholm (University of Arizona); Willard McCarty (University of Toronto).

**TR11 Drama and performance texts** Chair: Elli Mylonas* (Harvard University); John Lavagnino* (Brandeis University); Rosanne Potter (University of Iowa).

**TR12 Literary prose** Chair Thomas N. Corns* (University of Wales); Christian Delcourt (University of Liège).

**AI1 Linguistic description** Chair: D. Terence Langendoen* (University of Arizona); Stephen R. Anderson (Johns Hopkins University); Nicoletta Calzolari (University of Pisa); Geoffrey Sampson* (University of Sussex); Gary Simons* (Summer Institute of Linguistics).

**AI2 Spoken text** Chair: Stig Johansson* (University of Oslo); Jane Edwards (University of California at Berkeley); Andrew Rosta (University College London).

**AI3 Literary studies** Chair: Paul Fortier* (University of Manitoba); Christian Delcourt (University of Liège;); Ian Lancashire (University of Toronto); Rosanne Potter (University of Iowa); David Robey* (University of Manchester).

**AI4 Historical studies** Chair: Daniel Greenstein* (University of Glasgow); Peter Denley (Queen Mary Westfield College, London); Ingo Kropac (University of Graz); Hans Jørgen Marker (Danish Data Archive, Odense); Jan Oldervoll (University of Tromsø); Kevin Schurer (University of Cambridge); Donald Spaeth (Glasgow University); Manfred Thaller (Max-Planck-Institut für Geschichte, Göttingen).[144]

**AI5 Print dictionaries** Chairs: Robert Amsler* (Bell Communications Research) and Nicoletta Calzolari (University of Pisa); Susan Armstrong-Warwick (University of Geneva); John Fought (University of Pennsylvania); Louise Guthrie (University of New Mexico); Nancy M. Ide* (Vassar College); Frank Tompa (University of Waterloo); Carol Van Ess-Dykema (US Department of Defense); Jean Veronis (University of Aix-en-Provence).

**AI6 Machine lexica** Chair: Robert Ingria* (Bolt Beranek Newman Inc); Susan Armstrong-Warwick (University of Geneva); Nicoletta Calzolari (University of Pisa).

**AI7 Terminological data** Chair: Alan Melby* (Brigham Young University) Gerhard Budin (University of Vienna); Gregory Shreve (Kent State University); Richard Strehlow (Oak Ridge National Laboratory); Sue Ellen Wright (Kent State University).

## Advisory Board

Members of the TEI Advisory Board during the lifetime of the project are listed below, grouped under the name of the organization represented.

**American Anthropological Association**: Chad McDaniel (University of Maryland).

**American Historical Association**: Elizabeth A. R. Brown (Brooklyn College, CUNY).

**American Philological Association**: Jocelyn Penny Small (Rutgers University).

**American Philosophical Association**: Allen Renear (Brown University).

**American Society for Information Science**: Clifford A. Lynch (University of California).

**Association for Computing Machinery, Special Interest Group for Information Retrieval**: 1989–93: Scott Deerwester (University of Chicago); 1993- : Martha Evens (Illinois Institute of Technology).

**Association for Documentary Editing**: David Chesnutt (University of South Carolina).

**Association for History and Computing**: 1989–91: Manfred Thaller, Max-Planck-Institut für Geschichte, Göttingen; 1991- : Daniel Greenstein (Glasgow University).

**Association Internationale Bible et Informatique** 1989–93: Wilhelm Ott (University of Tübingen); 1993- : Winfried Bader (University of Tübingen).

**Canadian Linguistic Association**: Anne-Maria di Sciullo (Université du Québec à Montréal)

**Dictionary Society of North America**: Barbara Ann Kipfer (Independent Consultant).

**AAP Electronic Publishing Special Interest Group**: 1989–92: Betsy Kiser (OCLC); 1992- : Deborah Bendig and Andrea Keyhani (OCLC).

**International Federation of Library Associations and Institutions**: J. D. Byrum Jr. (The Library of Congress).

**Linguistic Society of America**: Stephen Anderson (The Johns Hopkins University)

**Modern Language Association**: Randall Jones (Brigham Young University) and Ian Lancashire (University of Toronto).

## Steering Committee Membership

Members of the Steering Committee of the TEI during the preparation of this work were:

**Association for Computational Linguistics**:

[144] This Workgroup was jointly sponsored by the Association for History and Computing.

- 1987–1993: Robert A. Amsler (Bell Communications Research);
- 1987–1993: Donald E. Walker (Bell Communications Research);
- 1993–1994: Susan Armstrong-Warwick (University of Geneva);
- 1994–1999: Judith Klavans (Columbia University).

**Association for Computers and the Humanities**:

- 1987–1999: Nancy M. Ide (Vassar College);
- 1987–1994: C. M. Sperberg-McQueen (University of Illinois at Chicago);
- 1994–1999: David Barnard (Queen's University).

**Association for Literary and Linguistic Computing**:

- 1987–1999: Susan M. Hockey (Center for Electronic Texts in the Humanities);
- 1987–1999: Antonio Zampolli (University of Pisa).

# I Colophon

The text of this manual was prepared electronically on a variety of systems. Most sections were originally drafted by members of the work groups and working committees of the TEI; all have been revised by the editors to achieve greater uniformity of style and greater consistency in the tag set.

The web release of the Guidelines was created using a library of XSLT stylesheets to convert to XHTML; the PDF version for printing was produced by conversion to LaTeX markup, processed using XeLaTeX. The XSLT libraries were written by Sebastian Rahtz.

Almost every available SGML and XML editor or processing program has been used at one time or another by the TEI; but without the open source implementations of XML parsers, editors and XSLT engines by James Clark, Richard Stallman, Michael Kay, and Daniel Veillard, the TEI could not survive, and we thank these individuals. We would also like to thank the staff at SyncRO Soft SRL, creators of the oXygen editor, for their support for the TEI during the creation on P5.

Many volunteers contributed to the preparation of this release of the Guidelines; we particularly note the work of Sabine Krott, Eva Radermacher and Arianna Ciula for their work in structuring the bibliographies.

The production and release process for TEI P5 was originally developed by Sebastian Rahtz for the TEI Technical Council.

# Index


Entries in italic font refer to examples, with those in bold italics referring to examples where the indexed element is the root.

<ab>, *171*, *226*, *227*, *437*, *525*, *556*, *592*, *601*, *602*, *610*, *613*, *620*, *721*, *834*, *884*, *908*, *1127*, *1605*, *2230*
<abbr>, *115*, *213*, *422–424*, *472*, *503*, *601*, *735*, *911*, *964*, *1127*, *1366*, *1370*
@absolute
  <when>, *288–290*, *311*, *605*, *2312*
<abstract>, *61*, *62*, *913*
<accMat>, *395*, *915*
<acquisition>, *361*, *396*, *531*, *533*, *918*, *1524*, *1661*, *1825*, *1826*
@active
  <interaction>, *569*, *1557*, *2299*
  <relation>, *512*, *513*, *523*, *524*, *1362*, *1677*, *2036*
<activity>, *571*, *920*, *2137*
<actor>, *251–254*, *922*, *1065*, *1068*, *1070*, *1902*
<add>, *102*, *103*, *171*, *429–435*, *440*, *445*, *448*, *451*, *453*, *459*, *460*, *892*, *925*, *1728*, *1738*, *2229*, *2230*, *2233*, *2354*
<additional>, *398*, *531*, *534*, *927*, *1826*
<additions>, *392*, *393*, *929*
<addName>, *492*, *494*, *932*, *1920*
<address>, *33*, *68*, *107*, *109*, *306*, *518*, *531*, *935*, *938*, *1362*, *1826*, *1886*, *1920*, *2285*
<addrLine>, *33*, *68*, *107*, *109*, *306*, *518*, *935*, *938*, *1362*, *1920*, *2285*
<addSpan>, *431*, *447*, *941*
@adj
  <node>, *692*, *693*, *1797–1799*
@adjFrom
  <node>, *694*, *1797–1799*
@adjTo
  <node>, *694*, *1797*, *1799*
<adminInfo>, *398*, *531*, *534*, *927*, *943*, *1826*
<affiliation>, *309*, *947*
<age>, *950*
@age
  <person>, *570*, *1667*, *1886*, *1909*, *1910*, *1913*
  <persona>, *1913*
@agent
  <damage>, *443*, *444*, *1188*
  <damageSpan>, *442*, *443*
  <gap>, *443*
@aloud
  <said>, *85*, *2092*, *2094*
<alt>, *453*, *525*, *624–626*, *953*, *957*
<alternate>, *751*, *752*, *769*, *793–795*, *799*, *955*, *977*, *1159*, *1332*, *1702*, *2129*, *2304*
<altGrp>, *525*, *625*, *626*, *957*
<altIdent>, *747*, *759*, *783*, *959*, *2388*
<altIdentifier>, *360*, *361*, *375*, *961*, *1781*
<am>, *379*, *423*, *482*, *601*, *964*
@ana
  <anchor>, *648*
  <c>, *313*, *650*, *674*
  <cl>, *650*
  <date>, *540*
  <join>, *648*
  <lg>, *51*, *2274*
  <m>, *650*
  <phr>, *650*, *674*
  <rdg>, *470*, *631*
  <s>, *648*
  <seg>, *647*, *648*
  , *2208*
  <w>, *649*, *650*, *674*
<analytic>, *145*, *147–149*, *151*, *159*, *160*, *165*, *966*, *1019*, *1647*, *1754*
<anchor>, *92*, *127*, *288–291*, *356*, *431–433*, *443*, *446*, *447*, *451*, *478*, *479*, *481*, *598*, *603*, *608*, *646*, *648*, *721*, *723*, *733*, *832*, *941*, *969*, *1088*, *1191*, *1227*, *2017*, *2183*
@anchored
  <note>, *584*, *1809*
<annotation>, *632*, *971*, *1640*
<annotationBlock>, *974*, *1640*
<anyElement>, *976*, *977*
<app>, *428*, *435*, *463–470*, *475–483*, *631*, *979*, *980*, *1590*, *1617*, *1644*, *2003*, *2428*, *2429*, *2433*
<appInfo>, *56*, *982*, *984*
<application>, *56*, *982*, *984*
<arc>, *691*, *693*, *695–697*, *699*, *702*, *986*, *1492*
<argument>, *186*, *187*, *988*
@arity
  <tree>, *705–708*, *2338*
@assertedValue
  <certainty>, *732*, *733*, *735*, *1087*, *1088*
@atLeast
  <gap>, *895*, *1453*
  <height>, *366*
  <measure>, *895*
  <num>, *2375*
  <width>, *366*, *737*, *738*
@atMost
  <gap>, *895*, *1453*
  <num>, *2375*
  <width>, *366*, *737*, *738*
<att>, *49*, *67*, *68*, *139*, *741*, *747*, *758*, *991*, *1156*, *1314*, *1367*, *1993*, *2023*, *2052*, *2285*
@attachment
  <surface>, *415*, *417*
<attDef>, *260*, *505*, *747*, *756*, *758*, *759*, *780–782*, *784*, *792*, *801*, *802*, *992*, *995*, *1321*, *1348*
<attList>, *260*, *505*, *747*, *756*, *772*, *780–782*, *784*, *792*, *801*, *995*, *1321*, *1348*

<attRef>, *772*, *997*
@atts
<specDesc>, *742*, *2187*, *2194*
<author>, *29*, *38*, *59*, *67*, *68*, *86*, *124*, *143–151*, *153*, *155*, *157–163*, *165*, *275*, *277*, *308*, *339*, *361*, *369*, *380*, *398*, *471*, *835*, *836*, *857*, *966*, *1000*, *1011*, *1014*, *1019*, *1043*, *1232*, *1647*, *1693*, *1754*, *1755*, *1767*, *1776*, *1853*, *2012*, *2033*, *2109*, *2127*, *2169*, *2285*
<authority>, *33*, *156*, *158*, *1002*
@autoPrefix
<content>, *753*
<availability>, *33*, *34*, *398*, *943*, *1004*, *1624*, *1979*, *2285*

<back>, *129*, *130*, *175*, *191*, *193*, *203*, *204*, *249*, *604*, *1006*, *1284*, *1644*
@baseForm
<m>, *641–643*, *1698*
@behaviour
<model>, *761–766*, *768*, *894*, *1741*, *1742*, *1745*, *1748*, *1876*, *1881*
<bibl>, *10*, *38*, *50*, *51*, *59*, *68*, *86*, *87*, *118*, *129*, *143–147*, *150*, *152*, *153*, *162*, *163*, *165*, *187*, *195*, *202*, *275*, *277*, *283*, *307*, *308*, *339*, *368*, *369*, *377*, *381*, *394*, *397–399*, *426*, *471*, *472*, *510*, *511*, *519*, *580*, *582–584*, *857*, *865*, *878*, *927*, *1011*, *1043*, *1103*, *1107*, *1116*, *1127*, *1232*, *1341*, *1562*, *1647*, *1693*, *1764*, *1776*, *1853*, *1910*, *1920*, *1945*, *2012*, *2109*, *2127*, *2169*, *2257*, *2273*, *2285*
<biblFull>, *144*, *306*, *307*, *1014*
<biblScope>, *10*, *35*, *145–152*, *154*, *159*, *160*, *163*, *165*, *368*, *377*, *397*, *878*, *966*, *1017*, *1019*, *1232*, *1693*, *1754*, *1755*, *1764*, *2130*, *2132*
<biblStruct>, *38*, *68*, *144*, *145*, *147–149*, *151*, *154–162*, *165*, *966*, *1019*, *1647*, *1754*, *1755*, *2033*
<bicond>, *686*, *687*, *1022*, *1540*
<binary>, *658*, *659*, *662*, *665*, *682*, *683*, *686*, *687*, *1022*, *1024*, *1144*, *1386*, *1388*, *1538*, *1540*, *2305*, *2399*, *2411*
<binaryObject>, *1026*
<binding>, *364*, *394*, *1029*
<bindingDesc>, *394*, *1029*, *1030*
<birth>, *370*, *488*, *507*, *525*, *570*, *1034*, *1910*
<bloc>, *498*, *517*, *1038*
<body>, *8*, *9*, *12*, *74*, *134*, *136*, *137*, *175*, *177*, *179*, *180*, *189*, *191*, *193*, *196*, *231*, *233*, *234*, *245*, *249*, *255*, *256*, *266*, *268*, *291*, *303*, *306*, *320*, *321*, *413*, *431*, *460*, *476*, *575*, *619*, *620*, *623*, *627*, *628*, *631*, *632*, *678*, *679*, *870*, *941*, *1040*, *1254*, *1405*, *1640*, *1644*, *1894*, *2208*, *2210*, *2280*, *2295*
@break
<lb>, *139*, *834*, *1610*
<pb>, *857*
<broadcast>, *277*, *1043*, *2012*
<byline>, *184*, *189*, *191*, *202*, *1045*, *1046*, *1287*, *2318*

<c>, *93*, *313*, *642*, *650*, *674*, *676*, *1048*
<caesura>, *236*, *1050*
<calendar>, *63*, *543*, *1052*, *1053*, *1055*
@calendar
<date>, *63*, *543*, *544*, *835*, *843*, *846*
<calendarDesc>, *63*, *543*, *1052*, *1053*, *1055*
<camera>, *270*, *271*, *1058*, *1060*, *2402*
<caption>, *271*, *1060*
<case>, *1063*
<castGroup>, *253*, *254*, *1065*, *1070*
<castItem>, *247–249*, *251–254*, *257*, *258*, *262*, *833*, *922*, *1065*, *1068*, *1070*, *1762*, *1902*, *1903*
<castList>, *247–249*, *251*, *252*, *254*, *257*, *258*, *262*, *1070*, *1762*, *1902*, *1903*
<catchwords>, *371*, *1074*, *1135*
<catDesc>, *51*, *52*, *382*, *487*, *849*, *851*, *1076*, *1078*, *1080*, *2273*, *2274*, *2298*
<category>, *51*, *52*, *382*, *487*, *849*, *851*, *1078*, *1080*, *2273*, *2274*, *2298*
<catRef>, *52*, *60*, *1080*, *2298*
@cause
<rdg>, *465*, *483*
<rdgGrp>, *468*
<redo>, *2017*
<retrace>, *449*, *450*
<cb>, *1082*, *1605*
<cell>, *100*, *224*, *546–549*, *1085*, *1127*, *2077*, *2264*
@cert
<abbr>, *911*
<certainty>, *898*, *1087*
<corr>, *11*, *97*, *426*, *441*, *876*
<event>, *512*
<ex>, *424*
<expan>, *424*
<geo>, *525*
<population>, *523*
<rdg>, *483*
<certainty>, *441*, *456*, *731–737*, *897*, *898*, *1087*, *1088*
<change>, *67*, *68*, *398*, *458–460*, *632*, *855*, *971*, *1092*, *1650*, *2010*, *2063*
@change
<add>, *451*, *459*, *2354*
<addSpan>, *447*
<del>, *451*, *2354*
<delSpan>, *447*
<line>, *459*, *460*
<metamark>, *447*
<mod>, *460*
<retrace>, *450*
<undo>, *451*, *2354*
<zone>, *447*
<channel>, *311*, *569*, *1076*, *1095*, *1965*, *2299*
<char>, *128*, *209*, *215*, *216*, *292*, *421*, *642*, *1096*, *1098*, *1687*, *2361*
<charDecl>, *209*, *213–215*, *292*, *421*, *1098*

@children
<iNode>, *705–708*, *1553*, *2338*
<root>, *705–708*, *1655*, *2075*, *2338*
<choice>, *11–13*, *97*, *99*, *106*, *115*, *279*, *294*, *316*, *422–428*, *437*, *440*, *441*, *503*, *537*, *592*, *600*, *735*, *876*, *911*, *964*, *980*, *1100*, *1165*, *1366*, *1370*, *1860*, *2027*, *2151*, *2210*, *2414*
<cit>, *86*, *187*, *195*, *202*, *322*, *334–337*, *339*, *341*, *342*, *344*, *347*, *1103*, *1341*, *1586*, *1851*, *1920*, *1945*, *2127*
<citeData>, *141*, *597*, *1105*
<citedRange>, *160*, *1107*
<citeStructure>, *140*, *141*, *596*, *597*, *1105*, *1110*
<cl>, *636–638*, *650*, *1113*
@class
<attRef>, *772*, *997*
<msContents>, *1764*
<msItem>, *381*, *382*, *1774*
<msItemStruct>, *1776*
<classCode>, *60*, *68*, *156*, *158*, *1115*, *2298*
<classDecl>, *68*, *382*, *849*, *851*, *1116*
<classes>, *749*, *768*, *772*, *781–783*, *792*, *795*, *803*, *1118*, *1123*, *1321*
<classRef>, *769*, *771*, *772*, *793–797*, *799*, *1120*, *1159*, *1702*
<classSpec>, *768*, *782*, *800*, *801*, *1123*
<climate>, *522*, *1127*
<closer>, *185*, *186*, *188*, *508*, *1131*, *1207*, *1950*, *2156*
<code>, *306*, *308*, *741*, *1133*, *1672*
@code
<occupation>, *1842*
<socecStatus>, *570*, *2162*
<collation>, *386*, *1135*
<collection>, *373*, *374*, *533*, *1137*, *1661*, *1781*, *1825*, *1831*
<colloc>, *332*, *1140*
<colophon>, *1142*
@cols
<table>, *100*, *546*, *547*, *2264*
@columns
<layout>, *361*, *388*, *1605*, *1607*, *1925*
@commodity
<measure>, *111*, *112*, *886*, *895*, *1711*, *1712*
<cond>, *686*, *687*, *1144*, *1432*, *2305*
<condition>, *387*, *394*, *1146*
<constitution>, *311*, *569*, *1076*, *1149*, *1965*, *2299*
<constraint>, *754–756*, *759*, *1150*, *1156*
<constraintDecl>, *754*, *1153*
<constraintSpec>, *754–756*, *759*, *1150*, *1156*
@contemporary
<binding>, *1029*
<seal>, *2113*
<content>, *751–753*, *769*, *772*, *774*, *779*, *793–797*, *799*, *803*, *955*, *976*, *977*, *1159*, *1193*, *1195*, *1198*, *1235*, *1321*, *1332*, *1702*, *2129*, *2304*
<conversion>, *54*, *55*, *859*, *860*, *1162*, *2377*, *2379*
@copyOf
<date>, *615*, *623*
<div2>, *623*
<eTree>, *711*, *713*, *715*, *1418*
<fs>, *667*
<l>, *483*, *615*
<seg>, *615*
<supplied>, *1325*
<corr>, *11–13*, *97*, *294*, *425–428*, *437*, *440*, *441*, *876*, *1100*, *1165*, *2151*, *2210*
<correction>, *43*, *68*, *427*, *575*, *576*, *1167*, *2285*
@corresp
<anchor>, *608*, *723*
<annotationBlock>, *1640*
<body>, *870*
<eLeaf>, *715*, *1418*
<eTree>, *715*, *1418*
<name>, *838*, *1943*
<person>, *871*
<place>, *871*
<ref>, *1944*
<seg>, *607*
<correspAction>, *64*, *65*, *1169*, *1172*
<correspContext>, *63*, *1170–1172*
<correspDesc>, *65*, *1172*
<country>, *109*, *306*, *309*, *373*, *374*, *496*, *498*, *517–519*, *525*, *530*, *531*, *533*, *534*, *935*, *1137*, *1176*, *1362*, *1661*, *1690*, *1825*, *1826*, *1831*, *1832*, *1910*, *1927*, *2210*
<creation>, *58*, *68*, *458–460*, *1092*, *1179*, *1650*
@cRef
<ptr>, *1977*
<ref>, *595*
<cRefPattern>, *139*, *140*, *594*, *595*, *1180*, *2022*, *2285*
@cssClass
<model>, *762*
<custEvent>, *399*, *531*, *534*, *1183*, *1185*, *1826*
<custodialHist>, *399*, *531*, *534*, *927*, *943*, *1185*, *1826*

<damage>, *442–444*, *1188*
<damageSpan>, *442*, *443*, *1191*
<dataFacet>, *757*, *1193*
<dataRef>, *505*, *757–759*, *780*, *781*, *784*, *801*, *992*, *995*, *1193*, *1195*, *1198*, *1200*, *1235*, *1321*
<dataSpec>, *1193*, *1195*, *1198*, *1233*, *1235*
<datatype>, *505*, *758*, *759*, *780*, *781*, *784*, *801*, *992*, *995*, *1193*, *1200*, *1321*
@datcat
<f>, *661*, *847*
<pos>, *353*, *848*, *849*
<symbol>, *663*, *684*, *848*, *850*
<taxonomy>, *849*
<date>, *10*, *31*, *33*, *34*, *38*, *58*, *63–65*, *68*, *113*, *114*, *144–162*, *165*, *186*, *188*, *251*, *252*, *275–277*, *306–310*, *312*, *349*, *350*, *352*, *361*, *385*, *396*, *398*, *399*, *459*, *487*, *489*, *500*, *511*, *540–544*, *570*, *571*, *614*, *615*, *623*,

*737*, *835*, *843–846*, *857*, *865*, *878*, *943*, *966*, *1011*, *1014*, *1019*, *1034*, *1043*, *1131*, *1169*, *1171*, *1172*, *1179*, *1183*, *1204*, *1205*, *1207*, *1232*, *1294*, *1300*, *1302*, *1362*, *1543*, *1647*, *1754*, *1755*, *1848*, *1902*, *1903*, *1920*, *1950*, *1954*, *1979*, *1982*, *1984*, *2012*, *2014*, *2033*, *2137*, *2169*, *2257*, *2285*, *2318*, *2461*, *2462*
<dateline>, *74*, *185*, *186*, *188*, *846*, *1131*, *1207*, *1848*, *1950*
@datingMethod
<date>, *543*, *544*, *843*, *844*
<event>, *843*
@datum
<geoDecl>, *1467*, *1469*
<death>, *488*, *525*, *1211*, *1910*
@decls
<div1>, *575*
, *245*
<geo>, *1356*, *1467*
<decoDesc>, *361*, *368*, *391*, *534*, *1212*, *1215*, *1925*
<decoNote>, *391*, *394*, *534*, *1215*
<def>, *322*, *326–328*, *333*, *334*, *336–339*, *344–347*, *349–352*, *356–358*, *539*, *1217*, *1237*, *1336*, *1339*, *1526*, *2006–2008*, *2127*, *2237*
<default>, *673*, *674*, *1219*
@default
<correction>, *575*, *576*
<editorialDecl>, *576*
<normalization>, *576*
<defaultVal>, *802*, *1220*
@defective
<explicit>, *379*, *381*
<incipit>, *379*, *381*, *1548*
<msItem>, *379*, *381*
<msItemStruct>, *1776*
@degree
<certainty>, *441*, *456*, *732–737*, *897*, *1088*
<node>, *691–693*, *1492*, *1799*
<purpose>, *569*, *1076*, *1965*, *1988*, *2299*
<del>, *103*, *286*, *294*, *430–435*, *445*, *446*, *448*, *451*, *453*, *456*, *460*, *895*, *1224*, *1453*, *1728*, *1738*, *2058*, *2229*, *2230*, *2233*, *2354*
@delim
<citeStructure>, *140*, *141*, *596*, *597*, *1110*
<refState>, *50*, *141*, *2025*
<delSpan>, *432*, *433*, *447*, *1227*
<depth>, *1229*, *1714*
<derivation>, *311*, *569*, *1076*, *1231*, *1232*, *1965*, *2299*
<desc>, *54*, *55*, *100*, *260*, *279–282*, *288*, *292*, *299*, *301*, *305*, *306*, *313*, *314*, *456*, *488*, *505*, *509–511*, *514*, *515*, *521–524*, *529*, *553*, *554*, *617*, *620*, *731*, *732*, *735*, *743*, *745–747*, *758*, *759*, *766–768*, *780*, *781*, *784*, *793*, *801–803*, *849*, *859*, *860*, *871*, *894*, *992*, *995*, *1087*, *1123*, *1156*, *1193*, *1195*, *1198*, *1233*, *1235*, *1321*, *1324*, *1348*, *1467*, *1546*, *1562*, *1572*, *1580*, *1652*, *1664*, *1711*, *1717*, *1741*, *1753*, *1805*, *1853*, *1883*, *1884*, *1910*, *1935*, *2052*, *2214*, *2293*, *2332*, *2374*, *2379*, *2388*, *2410*
<dictScrap>, *358*, *1237*
<dim>, *366*, *534*, *738*, *1240*, *1243*
<dimensions>, *361*, *366*, *385*, *1243*, *1925*
@direct
<said>, *2093*
@discrete
<sound>, *2165*
<distinct>, *82*, *1246*
<distributor>, *33*, *68*, *306*, *307*, *1249*, *2285*
<district>, *498*, *517*, *518*, *525*, *531*, *935*, *1252*, *1690*, *1826*
, *6*, *8*, *9*, *12*, *86*, *111*, *116*, *125*, *129*, *134*, *177*, *179*, *181*, *184–188*, *191*, *195*, *198–200*, *203*, *204*, *219*, *234*, *238*, *241*, *245*, *248*, *256*, *274*, *293*, *300*, *303*, *304*, *309*, *310*, *320*, *413*, *431*, *476*, *478*, *508*, *554*, *587*, *592*, *603*, *608–610*, *612*, *613*, *621–624*, *626*, *631*, *632*, *637*, *734*, *742*, *744*, *784*, *865*, *870*, *904*, *908*, *941*, *1006*, *1131*, *1207*, *1254*, *1284*, *1314*, *1405*, *1426*, *1510*, *1605*, *1631*, *1640*, *1644*, *1720*, *1950*, *2349*, *2402*
<div1>, *117*, *132*, *133*, *136*, *137*, *170*, *171*, *177–179*, *182*, *184*, *191*, *200*, *249*, *255*, *256*, *261*, *267*, *268*, *557*, *575*, *601*, *602*, *620*, *623*, *1258*, *1262*, *1284*, *1509*, *1635*
<div2>, *132*, *133*, *169–171*, *177–179*, *182*, *256*, *267*, *268*, *601*, *602*, *623*, *1258*, *1262*, *1265*, *1272*, *1277*, *1281*, *1509*
<div3>, *132*, *133*, *1262*, *1265*, *1269*, *1272*, *1277*, *1281*
<div4>, *1262*, *1269*, *1272*, *1277*, *1281*
<div5>, *1272*, *1277*, *1281*
<div6>, *1277*, *1281*
<div7>, *1281*
<divGen>, *129*, *130*, *198*, *1284*
<docAuthor>, *187*, *202*, *1046*, *1287*, *1920*
<docDate>, *46*, *202*, *1290*
<docEdition>, *1292*, *2318*
<docImprint>, *46*, *189*, *202*, *1290*, *1294*, *2318*
@docLang
<schemaSpec>, *55*
<docTitle>, *46*, *187*, *189*, *191*, *202*, *248*, *1287*, *1296*, *2295*, *2318*, *2321*
<domain>, *311*, *569*, *1076*, *1298*, *1965*, *2299*
@domains
<joinGrp>, *618*, *1575*
<linkGrp>, *585*, *604*, *606*, *609*
@dur
<date>, *540*, *542*, *2462*
<kinesic>, *1580*
<media>, *1717*
<pause>, *281*, *294*, *306*, *1892*
<recording>, *276*, *277*, *2012*, *2014*
<time>, *540*, *2462*
<vocal>, *2410*
@dur-iso
<date>, *845*, *2461*, *2462*

<time>, *2461, 2462*

@ed
<lb>, *857*, *1610*
<milestone>, *50*, *137, 138*, *889*, *1735*
<pb>, *1894*
<refState>, *50*, *141*
<edition>, *10*, *31*, *50*, *59*, *68*, *153, 154*, *307*, *878*, *1014*, *1019*, *1300*, *1301*, *2109*
<editionStmt>, *27*, *31*, *68*, *307*, *1014*, *1301, 1302*
<editor>, *68*, *124*, *145, 146*, *148*, *150*, *153*, *161, 162*, *1011*, *1092*, *1232*, *1305*
<editorialDecl>, *43*, *68*, *575, 576*, *1306*, *1802*, *2093, 2094*, *2285*
@edRef
<pb>, *857*
<education>, *570*, *1309, 1310*
<eg>, *1312*, *1367*
<egXML>, *1314, 1315*
<eLeaf>, *710, 711*, *713*, *715*, *717, 718*, *1317*, *1350*, *1418*, *2340*
<elementRef>, *11*, *751, 752*, *769*, *771*, *794*, *796*, *878*, *955*, *1159*, *1318*, *1332*, *2103*, *2129*, *2304*
<elementSpec>, *260*, *505*, *743*, *745–747*, *762*, *764–766*, *772*, *778–784*, *792–798*, *803*, *894*, *1321*, *1348*, *1741*, *1742*, *1745*, *1881*, *2189*
<ellipsis>, *100, 101*, *1324, 1325*
<email>, *107*, *309*, *508*, *1328*, *1920*
<emph>, *8, 9*, *80, 81*, *89*, *189*, *279*, *294*, *616*, *627*, *645*, *720, 721*, *724*, *726*, *735*, *1324*, *1331*, *1920*
<empty>, *1159*, *1332*
<encodingDesc>, *40*, *54, 55*, *68*, *139*, *245*, *572*, *576*, *678, 679*, *849*, *851*, *859*, *860*, *1334*, *1711*, *2285*, *2374*
@end
<annotationBlock>, *974*
<kinesic>, *306*
<post>, *306*
<u>, *289, 290*, *306*
@enjamb
<l>, *236*
<entry>, *320–322*, *324*, *326–328*, *331*, *333–339*, *342–345*, *347*, *349–353*, *358*, *849*, *1063*, *1103*, *1140*, *1217*, *1237*, *1336*, *1353*, *1459*, *1486*, *1489*, *1526*, *1529*, *1593*, *1613*, *1759*, *1819*, *1851*, *1900*, *1939*, *1972*, *2006–2008*, *2227*, *2237*, *2325*, *2445*
<entryFree>, *357*, *1082*, *1339*
@eol
<hyphenation>, *68*, *1531*
<epigraph>, *86*, *187*, *191*, *202*, *204*, *1341*, *1426*
<epilogue>, *249*, *1343*
<equipment>, *276, 277*, *1346*, *2012*, *2014*
<equiv>, *746, 747*, *759*, *798*, *801*, *894*, *1348*, *2388*
<eTree>, *710, 711*, *713*, *715*, *717, 718*, *1350*, *1418*
<etym>, *337, 338*, *345*, *349–352*, *357, 358*, *537*, *1237*, *1339*, *1353*, *1593*, *1658*, *2445*
@evaluate
<joinGrp>, *620*
<link>, *586*
<event>, *488*, *511, 512*, *521*, *524, 525*, *529*, *843*, *1356, 1357*, *1361, 1362*, *1652*
<eventName>, *500*, *1356*, *1361, 1362*
@evidence
<gender>, *507*
<origin>, *1867*
<persPronouns>, *507*
<ex>, *101*, *361*, *379*, *381*, *393*, *395*, *413*, *423, 424*, *482*, *964*, *1366*, *1370*, *1774*, *1776*, *2085*
@except
<anyElement>, *976*
<classRef>, *772*
<moduleRef>, *771*, *1750*
@exclude
<app>, *483*
<arc>, *702*
<date>, *623*
<div2>, *623*
<eTree>, *711*, *713*
<name>, *623*
<seg>, *622*, *625*
<u>, *621–623*
<exemplum>, *1367*
<expan>, *115*, *379*, *422–424*, *503*, *735*, *911*, *964*, *1366*, *1370*
@expand
<pos>, *881*
<explicit>, *369*, *379–381*, *1373*, *1693*, *1774*
<extent>, *27*, *32*, *144*, *147*, *153*, *157*, *158*, *162*, *165*, *307*, *361*, *385, 386*, *1014*, *1375*, *1925*, *2033*
@extent
<damageSpan>, *442*
<dim>, *1240*
<gap>, *100*, *181*, *286*, *293*, *445, 446*, *592*, *854*, *895*, *1453*, *1774*
<height>, *854*
<orth>, *349, 350*, *354*, *356*
<pron>, *339*, *348–350*, *356*, *1972*, *2006*
<width>, *366*

<f>, *314*, *541*, *658–675*, *683*, *686, 687*, *847, 848*, *1022*, *1024*, *1144*, *1219*, *1376, 1377*, *1402*, *1430*, *1538*, *1540*, *2208*, *2223*, *2261*, *2305*, *2393*, *2397*, *2404*, *2407*
@facs
<figure>, *413*
<head>, *413*
<locus>, *368*, *1693*
<p>, *413*
<pb>, *404*, *413*, *1894*
<surface>, *2171*
<w>, *413*

<facsimile>, *405–408*, *410*, *411*, *413*, *419*, *1378*, *1495*, *1496*, *2247*, *2280*, *2448*, *2473*
<factuality>, *311*, *569*, *1076*, *1381*, *1965*, *2299*
<faith>, *1384*
<fDecl>, *680*, *682–684*, *687*, *850*, *1386*, *1388*, *1434*, *1436*, *1438*, *2399*, *2411*
<fDescr>, *682–684*, *687*, *850*, *1386*, *1388*, *1438*, *2399*, *2411*
@feats
  <fs>, *663*, *667*, *675*
@feature
  <shift>, *284*, *2148*
<figDesc>, *555*, *556*, *587*, *1390*, *1393*, *1495*, *1717*, *2318*, *2472*
<figure>, *118*, *213*, *299*, *301*, *302*, *305*, *313*, *314*, *413*, *554–556*, *587*, *1390*, *1393*, *1483*, *1495*, *1717*, *2318*, *2472*
<fileDesc>, *22*, *27*, *68*, *572*, *632*, *678*, *679*, *687*, *1394*, *2280*, *2285*
<filiation>, *381*, *1397*
@filter
  <equiv>, *747*
<finalRubric>, *1400*, *1776*
<fLib>, *662*, *665*, *675*, *1402*
@flipping
  <surface>, *415*, *417*
<floatingText>, *196*, *266*, *1405*
<floruit>, *510*, *1408*
<foliation>, *386*, *387*, *1410*
@follow
  <iNode>, *708*, *1553*
@force
  <pc>, *77*, *643*
<foreign>, *55*, *61*, *79*, *80*, *89*, *294*, *396*, *865*, *1413*, *2051*
<forename>, *13*, *144*, *145*, *147–151*, *155*, *159*, *165*, *370*, *490–494*, *507*, *511*, *537*, *932*, *1011*, *1019*, *1362*, *1416*, *1465*, *1755*, *1791*, *1886*, *1907*, *1920*, *2073*
<forest>, *715*, *1418*, *1655*
<form>, *320–322*, *326–328*, *330–335*, *337*, *341–345*, *347–357*, *536–539*, *849*, *1063*, *1082*, *1103*, *1140*, *1217*, *1336*, *1353*, *1421*, *1459*, *1486*, *1489*, *1526*, *1529*, *1569*, *1593*, *1613*, *1658*, *1759*, *1819*, *1823*, *1851*, *1873*, *1900*, *1939*, *1972*, *2006–2008*, *2221*, *2227*, *2237*, *2260*, *2325*, *2383*, *2445*
@form
  <objectDesc>, *361*, *384*, *385*, *400*, *1781*, *1830*, *1925*, *2243*
<formula>, *386*, *550*, *552*, *1135*, *1424*
@formula
  <conversion>, *54*, *55*, *859*, *860*, *1162*, *2377*, *2379*
@from
  <affiliation>, *947*
  <app>, *478*, *479*, *481*, *980*, *1644*
  <arc>, *691*, *693*, *695–697*, *699*, *702*, *986*, *1492*
  <biblScope>, *145*, *147*, *148*, *159*, *160*, *165*, *1017*, *1754*
  <citedRange>, *1107*
  <conversion>, *54*, *55*, *1162*
  <custEvent>, *534*
  <date>, *114*, *489*, *542*, *846*, *865*, *1172*
  <education>, *1309*
  <event>, *488*, *1356*, *1361*, *1362*, *1652*
  <locus>, *367*, *368*, *379*, *381*, *388*, *389*, *1607*, *1693*, *1696*, *1776*
  <orgName>, *514*
  <persName>, *154*
  <residence>, *488*, *489*, *738*
  <rt>, *92*
  , *644–646*, *655*, *1640*, *2180*, *2183*
  <state>, *510*, *514*, *521*, *2214*
@from-custom
  <event>, *843*
@fromUnit
  <conversion>, *54*, *55*, *859*, *860*, *1162*, *2377*, *2379*
<front>, *134*, *175*, *189*, *191*, *193*, *194*, *200*, *202*, *231*, *247–249*, *266*, *1284*, *1315*, *1426*, *1498*, *2135*, *2295*
<fs>, *314*, *541*, *658–661*, *663–668*, *670–673*, *675*, *678*, *679*, *683*, *684*, *686*, *687*, *847*, *848*, *1022*, *1144*, *1377*, *1430*, *1432*, *1538*, *1540*, *2208*, *2305*, *2397*, *2404*
<fsConstraints>, *680*, *687*, *1432*, *1436*
<fsdDecl>, *678*, *679*, *687*, *1434*
<fsDecl>, *678–680*, *687*, *1434*, *1436*, *1438*
<fsDescr>, *680*, *687*, *1436*, *1438*
<fsdLink>, *679*, *1434*, *1440*
@full
  <forename>, *144*, *148*, *159*, *491*, *494*
  <genName>, *494*
  <roleName>, *493*
@function
  <cl>, *638*, *1113*
  <material>, *361*, *364*
  <metamark>, *100*, *101*, *415*, *417*, *447*, *448*, *452*, *1324*, *1728*
  <phr>, *638*, *723*, *1923*
  <w>, *723*
<funder>, *30*, *1442*
@fVal
  <f>, *661*, *663*, *665*, *667*, *848*
<fvLib>, *663*, *665*, *675*, *676*, *848*, *1445*
<fw>, *413*, *457*, *1448*

<g>, *128*, *213*, *215*, *216*, *293*, *379*, *421–423*, *467*, *468*, *482*, *601*, *642*, *964*, *979*, *1450*, *1536*, *1590*, *2429*
<gap>, *99*, *100*, *181*, *286*, *293*, *294*, *317*, *318*, *361*, *381*, *436*, *437*, *443–446*, *592*, *838*, *854*, *895*, *898*, *1224*, *1373*, *1453*, *1774*, *1944*, *2230*
<gb>, *1456*
<gen>, *333–335*, *1063*, *1459*, *2007*, *2008*
<gender>, *507*, *508*, *1462*
@gender
  <person>, *1910*
@generatedBy

<c>, *313*
<figure>, *299*, *301*, *302*, *305*, *313*, *314*
<post>, *299*, *301*, *302*, *305*, *306*, *313–318*, *838*, *1943*, *1944*
<seg>, *1945*
<signed>, *301*, *309–312*, *838*, *1944*, *1945*
<time>, *302*, *310*, *314*, *838*, *1944*, *1945*
<genName>, *492–494*, *1465*
<geo>, *502*, *517*, *519*, *521*, *525*, *531*, *1356*, *1361*, *1362*, *1467*, *1689*, *1826*, *2210*
<geoDecl>, *1467*, *1469*
<geogFeat>, *499*, *539*, *1473*, *1477*, *1846*
<geogName>, *498*, *499*, *539*, *1473*, *1477*, *1846*, *1932*
<gi>, *43*, *49*, *67*, *68*, *139*, *589*, *741*, *742*, *747*, *758*, *1123*, *1312*, *1314*, *1362*, *1367*, *1479*, *1672*, *1802*, *1959*, *1986*, *1993*, *2023*, *2052*, *2124*, *2285*, *2387*, *2388*
@gi
<tagUsage>, *47*, *260*, *849*, *1792*, *2269*, *2271*
@given
<certainty>, *732*, *733*, *1088*
<gloss>, *88*, *89*, *121*, *337*, *338*, *345*, *349*, *350*, *352*, *537*, *745*, *747*, *758*, *759*, *798*, *802*, *851*, *1321*, *1353*, *1482*, *1725*, *2289*, *2388*, *2392*
<glyph>, *209*, *213–216*, *1450*, *1483*
@gradual
<writing>, *283*
<gram>, *322*, *332*, *337*, *1486*
<gramGrp>, *322*, *326–328*, *330–335*, *337*, *345*, *349*, *350*, *352*, *353*, *356*, *357*, *848*, *849*, *881*, *882*, *1063*, *1140*, *1217*, *1336*, *1459*, *1486*, *1489*, *1526*, *1759*, *1819*, *1900*, *1939*, *1972*, *2006–2008*, *2227*, *2237*, *2325*
<graph>, *691–697*, *699*, *702*, *1492*, *1797–1799*
<graphic>, *118*, *131*, *213*, *302*, *313*, *314*, *405–408*, *410*, *411*, *413*, *414*, *418*, *419*, *554–556*, *587*, *1378*, *1390*, *1393*, *1483*, *1495*, *1496*, *1805*, *1889*, *2171*, *2247*, *2249*, *2280*, *2448*, *2472*, *2473*
<group>, *175*, *189*, *191*, *193*, *194*, *870*, *1315*, *1498*, *2295*
@group
<damage>, *443*, *1188*
<damageSpan>, *443*

@hand
<add>, *429–431*, *440*
<addSpan>, *431*, *941*
<del>, *430*, *431*, *456*, *2229*
<mod>, *2017*
<rdg>, *465*, *469*
<redo>, *2017*
<restore>, *435*, *2058*
<retrace>, *460*
<rt>, *94*
<handDesc>, *361*, *389*, *390*, *1500*, *1501*, *1925*
<handNote>, *389*, *390*, *429–431*, *439*, *897*, *941*, *1500*, *1503*, *1505*
<handNotes>, *439*, *1503*, *1505*
@hands
<handDesc>, *389*, *1501*
<handShift>, *439*, *460*, *1507*
<head>, *86*, *87*, *100*, *120*, *121*, *129–131*, *134*, *145*, *146*, *169–171*, *178*, *182*, *184–187*, *189*, *191*, *193*, *195*, *196*, *199*, *200*, *203*, *204*, *231*, *248*, *249*, *251*, *253–256*, *259*, *264*, *265*, *300*, *320*, *376*, *413*, *472*, *476*, *478*, *525*, *543*, *546–549*, *554–556*, *587*, *601*, *602*, *612*, *613*, *617*, *628*, *631*, *742*, *744*, *843*, *865*, *870*, *873*, *874*, *904*, *1006*, *1070*, *1127*, *1254*, *1258*, *1262*, *1265*, *1269*, *1272*, *1277*, *1281*, *1284*, *1314*, *1343*, *1390*, *1393*, *1426*, *1495*, *1509*, *1510*, *1566*, *1586*, *1635*, *1644*, *1647*, *1652*, *1664*, *1712*, *1717*, *1720*, *1781*, *2091*, *2135*, *2264*, *2318*, *2349*, *2472*
<headItem>, *122*, *1512*, *1514*, *1586*
<headLabel>, *122*, *1512*, *1514*, *1586*
<height>, *361*, *366*, *385*, *854*, *1243*, *1517*, *1714*, *1925*
@height
<graphic>, *2472*
<heraldry>, *372*, *1519*
<hi>, *46*, *81*, *89*, *196*, *232*, *238*, *270*, *344*, *424*, *456*, *470*, *472*, *496*, *543*, *601*, *762*, *873*, *874*, *1082*, *1424*, *1510*, *1522*, *1817*, *2203*, *2321*, *2328*, *2402*
<history>, *361*, *396*, *531*, *533*, *534*, *1524*, *1661*, *1825*, *1826*
<hom>, *324*, *327*, *345*, *1526*
<hyph>, *326–328*, *330*, *357*, *1217*, *1529*, *2237*, *2260*
<hyphenation>, *68*, *1531*

<ident>, *43*, *741*, *1533*
@ident
<application>, *56*, *982*, *984*
<attDef>, *260*, *505*, *747*, *756*, *758*, *759*, *780–782*, *784*, *792*, *801*, *802*, *992*, *995*, *1321*, *1348*
<classSpec>, *768*, *782*, *800*, *801*, *1123*
<constraintSpec>, *754–756*, *759*, *1150*, *1156*
<dataSpec>, *1193*, *1195*, *1198*, *1233*, *1235*
<elementSpec>, *260*, *505*, *743*, *745–747*, *762*, *764–766*, *772*, *778–784*, *792–798*, *803*, *894*, *1321*, *1348*, *1741*, *1742*, *1745*, *1881*, *2189*
<language>, *58*, *68*, *897*, *1601*, *1602*, *1965*
<macroSpec>, *799*, *1702*
<moduleSpec>, *743*, *1753*
<paramSpec>, *766*, *767*, *1883*, *1884*
<prefixDef>, *589*, *590*, *849*, *1672*, *1959*
<schemaSpec>, *3*, *4*, *55*, *260*, *770–773*, *778*, *785*, *798*, *878*, *1120*, *1195*, *1700*, *2103*
<transcriptionDesc>, *277*, *2334*
<valItem>, *260*, *505*, *745*, *759*, *766*, *767*, *781*, *801*, *802*, *1348*, *1883*, *1884*, *2388*, *2392*
<idno>, *33*, *35*, *144*, *149*, *153*, *155–158*, *307*, *308*, *360*, *361*, *373–375*, *399–401*, *428*, *502*, *519*, *531*, *533*, *631*, *632*, *743*, *961*, *1137*, *1356*, *1361*, *1536*, *1555*, *1661*,

*1753*, *1767*, *1770*, *1772*, *1781*, *1825*, *1826*, *1831*, *2043*, *2109*, *2132*, *2257*, *2285*
<if>, *683*, *687*, *1538*
<iff>, *686*, *687*, *1022*, *1540*
<imprimatur>, *202*, *1542*
<imprint>, *38*, *68*, *144*, *145*, *147–149*, *151*, *154–162*, *165*, *966*, *1019*, *1543*, *1647*, *1754*, *1755*, *1982*, *2033*
<incident>, *279*, *280*, *282*, *1546*
<incipit>, *361*, *369*, *379–381*, *1548*, *1693*, *1774*
@include
<classRef>, *771*
<moduleRef>, *771*, *1120*, *1750*
@inDegree
<node>, *693*, *694*, *696*, *697*, *699*, *1799*, *1800*
@indentLevel
<post>, *300*, *301*, *309*, *310*, *312*, *838*, *1944*
<index>, *107*, *127–129*, *899*, *1551*
@indexName
<index>, *128*, *899*, *1551*
<iNode>, *705–708*, *1553*, *2338*
@instant
<del>, *453*
<institution>, *375*, *531*, *533*, *1555*, *1661*, *1825*, *1826*, *1831*, *2043*
<interaction>, *311*, *569*, *1076*, *1557*, *1965*, *2299*
<interp>, *647*, *648*, *650*, *651*, *1560*, *1562*
<interpGrp>, *648*, *650*, *651*, *1562*
<interpretation>, *43*, *68*, *1564*
@interval
<when>, *288–290*, *605*, *2312*, *2420*
<item>, *6*, *62*, *68*, *85*, *111*, *117*, *119–122*, *125*, *126*, *139*, *146*, *184*, *187*, *199*, *200*, *203*, *204*, *221*, *260*, *386*, *391*, *393*, *427*, *570*, *587*, *617*, *739*, *1006*, *1272*, *1277*, *1281*, *1314*, *1510*, *1512*, *1514*, *1566*, *1572*, *1586*, *1634–1636*, *1720*, *1781*, *1993*, *2051*, *2135*, *2156*, *2185*, *2285*, *2298*, *2347*
@iterated
<kinesic>, *288*, *1580*
<vocal>, *2410*
<iType>, *331*, *1569*

<join>, *237*, *268*, *617*, *618*, *620*, *648*, *726*, *1572*, *1575*
@join
<pc>, *654*, *885*, *1897*
<w>, *316*, *654*, *885*
<joinGrp>, *618*, *620*, *1575*

@key
<bibl>, *144*
<classRef>, *769*, *771*, *772*, *793–797*, *799*, *1120*, *1159*, *1702*
<country>, *517*, *935*, *1176*, *1690*, *1910*, *2210*
<dataRef>, *505*, *758*, *759*, *780*, *781*, *784*, *801*, *992*, *995*, *1195*, *1200*, *1321*
<elementRef>, *11*, *751*, *752*, *769*, *771*, *794*, *796*, *878*, *955*, *1159*, *1318*, *1332*, *2103*, *2129*, *2304*
<event>, *524*
<macroRef>, *772*, *779*, *803*, *1700*
<memberOf>, *749*, *768*, *772*, *781–783*, *792*, *795*, *803*, *1118*, *1123*, *1321*, *1722*
<moduleRef>, *3*, *4*, *55*, *260*, *770–773*, *778*, *785*, *1120*, *1750*, *2103*
<name>, *62*, *105*, *361*, *486*, *589*, *835*, *836*
<nationality>, *510*, *1795*
<orgName>, *1000*, *1858*
<origPlace>, *1781*
<placeName>, *521*, *524*, *1846*, *1858*
<region>, *2210*
<rs>, *105*, *106*
<settlement>, *400*
<specDesc>, *742*, *2186*, *2187*, *2194*
<trait>, *509*
<keywords>, *59*, *68*, *1116*, *1577*, *2298*
<kinesic>, *279*, *288*, *306*, *1580*

<l>, *6*, *12*, *51*, *81*, *86*, *101–103*, *123*, *124*, *132*, *133*, *137*, *167–171*, *191*, *195*, *196*, *219*, *227*, *231–243*, *245*, *249*, *250*, *257–259*, *263–268*, *432–434*, *437*, *439*, *443*, *460*, *465*, *466*, *469*, *476–481*, *483*, *580*, *582–584*, *615*, *618*, *626*, *628*, *637*, *638*, *720–722*, *724–726*, *728*, *833*, *857*, *876*, *886*, *904*, *1040*, *1050*, *1113*, *1188*, *1224*, *1324*, *1325*, *1341*, *1343*, *1405*, *1507*, *1572*, *1583*, *1610*, *1620*, *1640*, *1644*, *1860*, *1969*, *2066*, *2119*, *2196*, *2203*, *2274*, *2295*, *2328*, *2414*, *2436*
<label>, *54–56*, *111*, *119–122*, *124*, *139*, *188*, *200*, *221*, *260*, *427*, *488*, *509–511*, *522*, *524*, *534*, *691–697*, *699*, *702*, *705–708*, *710*, *711*, *713*, *715*, *717*, *718*, *859*, *860*, *982*, *984*, *986*, *1317*, *1325*, *1350*, *1357*, *1418*, *1492*, *1512*, *1514*, *1553*, *1586*, *1587*, *1614*, *1652*, *1711*, *1797–1800*, *1950*, *1993*, *2075*, *2135*, *2214*, *2332*, *2338*, *2340*, *2374*, *2377*, *2379*
@label
<rhyme>, *242*, *2066*
<lacunaEnd>, *475*, *1589*
<lacunaStart>, *1590*
<lang>, *337*, *338*, *345*, *349–352*, *357*, *358*, *537*, *1237*, *1339*, *1353*, *1593*
@lang
<code>, *1133*
<langKnowledge>, *507*, *570*, *1595*, *1596*, *1598*
<langKnown>, *507*, *570*, *1596*, *1598*
<language>, *58*, *68*, *897*, *1601*, *1602*, *1965*
<langUsage>, *58*, *68*, *897*, *1601*, *1602*, *1965*
<layout>, *361*, *388*, *1605*, *1607*, *1830*, *1925*
<layoutDesc>, *361*, *388*, *1607*, *1830*, *1925*
<lb>, *46*, *101*, *112*, *138*, *139*, *167*, *188*, *202*, *224*, *227*, *364*, *369*, *381*, *393*, *413*, *437*, *508*, *543*, *592*, *601*, *643*, *721*, *834*,

*857*, *873*, *874*, *1510*, *1587*, *1610*, *1693*, *1776*, *2051*, *2203*, *2206*, *2210*, *2229*, *2328*
<lbl>, *322*, *331*, *343*, *344*, *349*, *350*, *1613*
<leaf>, *705–708*, *1614*, *2075*, *2338*
<lem>, *464–470*, *475*, *477–479*, *481*, *483*, *631*, *979*, *1590*, *1617*, *2003*, *2428*, *2429*, *2433*
@lemma
<pc>, *654*
<w>, *316–318*, *641*, *643*, *654*, *655*, *882*, *883*, *2414*
@lemmaRef
<w>, *641*, *883*
@length
<refState>, *50*, *2025*
@level
<langKnown>, *507*, *570*, *1596*, *1598*
<title>, *10*, *35*, *144–154*, *157–163*, *165*, *398*, *865*, *878*, *966*, *1011*, *1019*, *1232*, *1754*, *1755*, *1910*, *2130*, *2169*
<lg>, *12*, *51*, *101*, *123*, *124*, *137*, *168*, *169*, *191*, *219*, *227*, *231–234*, *237–239*, *241–243*, *245*, *249*, *259*, *265–268*, *720*, *722*, *724*, *726*, *876*, *904*, *1324*, *1325*, *1343*, *1620*, *1640*, *1644*, *2066*, *2274*
<licence>, *34*, *632*, *971*, *1004*, *1623*, *1624*
<line>, *414*, *415*, *417*, *446*, *448*, *452*, *459*, *460*, *1626*, *1728*, *2017*, *2171*, *2354*
<link>, *243*, *289*, *291*, *557*, *580*, *583–586*, *604*, *606–609*, *613*, *614*, *617*, *624*, *648*, *652*, *676*, *1629*, *1631*
<linkGrp>, *243*, *289*, *291*, *557*, *585*, *604*, *606–609*, *613*, *614*, *648*, *652*, *676*, *893*, *1629*, *1631*
<list>, *6*, *62*, *68*, *85*, *111*, *117*, *119–122*, *125*, *126*, *139*, *146*, *184*, *187*, *199*, *200*, *203*, *221*, *260*, *386*, *391*, *393*, *427*, *570*, *587*, *617*, *739*, *1006*, *1272*, *1277*, *1281*, *1314*, *1510*, *1512*, *1514*, *1566*, *1572*, *1586*, *1634–1636*, *1720*, *1781*, *1993*, *2051*, *2135*, *2156*, *2185*, *2285*, *2298*, *2347*
<listAnnotation>, *632*, *1640*
<listApp>, *1644*
<listBibl>, *129*, *145*, *398*, *472*, *857*, *865*, *927*, *1647*, *1910*
<listChange>, *458–460*, *1092*, *1650*
<listEvent>, *524*, *525*, *1356*, *1361*, *1652*
<listForest>, *1655*
<listNym>, *536*, *1658*
<listObject>, *531*, *533*, *1661*, *1825*, *1826*
<listOrg>, *515*, *1664*
<listPerson>, *85*, *279*, *282*, *287*, *293*, *309*, *311*, *430*, *503*, *525*, *1356*, *1362*, *1667*, *1677*, *1886*, *1943*, *2148*
<listPlace>, *520*, *521*, *523*, *525*, *631*, *632*, *1361*, *1670*, *1677*, *2210*
<listPrefixDef>, *589*, *849*, *1672*
<listRef>, *748*, *1675*, *1910*
<listRelation>, *513*, *1362*, *1667*, *1677*, *1678*, *1886*
<listTranspose>, *452*, *1680*
<listWit>, *470–472*, *1683*, *2431*
@loc
<app>, *476*, *980*
<locale>, *571*, *1685*, *2137*
<localProp>, *209*, *213*, *216*, *1096*, *1483*, *1687*, *2361*
<location>, *502*, *517–519*, *521*, *525*, *531*, *1356*, *1361*, *1362*, *1467*, *1689*, *1690*, *1826*, *2210*
@location
<pos>, *356*
<variantEncoding>, *476*, *478*, *479*, *2395*
<locus>, *367–369*, *377–381*, *386–389*, *391*, *393*, *1607*, *1693*, *1696*, *1764*, *1774*, *1776*, *2085*
@locus
<certainty>, *441*, *456*, *731–737*, *897*, *898*, *1087*, *1088*
<respons>, *441*, *739*, *2051*, *2052*
<locusGrp>, *368*, *369*, *1696*
@lrx
<surface>, *407*, *408*, *410*, *411*, *414*, *417–419*, *1378*, *1889*, *2247*, *2249*, *2448*, *2473*
<zone>, *407*, *408*, *410*, *411*, *414*, *417*, *419*, *2448*
@lry
<surface>, *407*, *408*, *410*, *411*, *414*, *417–419*, *1378*, *1889*, *2247*, *2249*, *2448*, *2473*
<zone>, *407*, *408*, *410*, *411*, *414*, *417*, *419*, *2448*

<m>, *641–643*, *650*, *883*, *1698*
<macroRef>, *772*, *779*, *803*, *1700*
<macroSpec>, *799*, *1702*
@mainLang
<textLang>, *163*, *164*, *361*, *382*, *383*, *400*, *1774*, *1776*, *1781*, *2302*
<mapping>, *209*, *214–216*, *1096*, *1098*, *1687*, *1705*, *2361*
@marks
<punctuation>, *1986*
<quotation>, *43*, *68*, *83*, *1306*, *1993*, *1994*, *2093*, *2094*
@match
<certainty>, *734*, *736*, *737*, *898*
<citeStructure>, *140*, *141*, *596*, *597*, *1105*, *1110*
<precision>, *737*, *738*, *1954*
<respons>, *739*, *2051*, *2052*
@matchPattern
<cRefPattern>, *139*, *140*, *594*, *595*, *1180*, *2022*, *2285*
<prefixDef>, *589*, *590*, *849*, *1672*, *1959*
<material>, *361*, *364*, *385*, *394*, *533*, *534*, *1661*, *1708*, *1825*, *1830*, *2243*, *2244*, *2418*
@material
<supportDesc>, *361*, *1830*, *1925*
@max
<height>, *366*
<measure>, *112*, *1712*
<numeric>, *660*, *669*, *1821*
@maxOccurs
<alternate>, *751*, *793–795*, *799*, *977*, *1332*, *1702*, *2304*
<anyElement>, *976*
<classRef>, *793*, *795–797*, *1159*
<datatype>, *758*, *995*, *1200*
<elementRef>, *751*, *752*, *1159*, *2129*

<sequence>, *795–797*
<measure>, *32*, *111–113*, *307*, *386*, *500*, *518*, *525*, *886*, *895*, *1375*, *1711*, *1712*, *1932*, *2374*, *2375*
<measureGrp>, *112*, *1714*
<media>, *1103*, *1717*
@medium
<handNote>, *390*, *439*, *1503*, *1505*
<handShift>, *439*, *1507*
<meeting>, *154*, *1720*
<memberOf>, *749*, *768*, *772*, *781–783*, *792*, *795*, *803*, *1118*, *1123*, *1321*, *1722*
<mentioned>, *80*, *87*, *88*, *124*, *195*, *337–339*, *345*, *357*, *364*, *466*, *469*, *530*, *537*, *640*, *731*, *1353*, *1593*, *1658*, *1725*, *1808*, *2445*, *2472*
@mergedIn
<orth>, *355*, *2007*, *2008*
<pron>, *355*
@met
, *238*, *241*
<l>, *239*, *240*, *1583*
<lg>, *239*, *241*
<seg>, *240*
<metamark>, *100*, *101*, *415*, *417*, *447*, *448*, *452*, *1324*, *1325*, *1728*
<metDecl>, *241*, *244*, *245*, *1730*
@method
<correction>, *68*
<normalization>, *1802*
<variantEncoding>, *476*, *478*, *479*, *2395*
<metSym>, *241*, *244*, *245*, *1730*, *1733*
<milestone>, *50*, *136–138*, *889*, *1735*
@mimeType
<binaryObject>, *1026*
<equiv>, *747*
<graphic>, *131*
<media>, *1103*, *1717*
<ptr>, *553*, *554*
<ref>, *119*, *880*
@min
<height>, *366*
<measure>, *112*
@minOccurs
<alternate>, *751*, *793–795*, *799*, *977*, *1332*, *1702*, *2304*
<anyElement>, *976*
<classRef>, *793*, *795–797*
<datatype>, *758*, *1200*
<elementRef>, *751*, *752*, *1159*, *2129*
<sequence>, *795–797*
<mod>, *446*, *453*, *460*, *1738*, *2017*
@modality
<post>, *299–302*, *305*, *306*, *309–314*, *317*, *318*
@mode
<alt>, *453*, *624–626*, *953*
<altGrp>, *626*, *957*
<attDef>, *260*, *505*, *747*, *780–782*, *784*, *792*, *1348*
<channel>, *569*, *1076*, *1095*, *1965*, *2299*
<classSpec>, *782*
<classes>, *781*, *782*
<elementSpec>, *260*, *505*, *747*, *762*, *764–766*, *772*, *778–784*, *793*, *894*, *1348*, *1741*, *1742*, *1745*, *1881*
<memberOf>, *782*
<model>, *761–766*, *768*, *894*, *1741*, *1742*, *1745*, *1748*, *1876*, *1881*
<modelGrp>, *763*, *765*, *1745*, *1748*
<modelSequence>, *766*, *1745*, *1748*
@module
<classSpec>, *800*, *801*, *1123*
<dataSpec>, *1233*, *1235*
<elementSpec>, *505*, *745–747*, *779–783*, *795*, *798*, *803*, *1321*
<macroSpec>, *799*, *1702*
<moduleRef>, *3*, *4*, *55*, *260*, *770–774*, *778*, *785*, *1120*, *1750*, *2103*
<moduleSpec>, *743*, *1753*
<monogr>, *38*, *68*, *144*, *145*, *147–149*, *151*, *154–162*, *165*, *966*, *1019*, *1647*, *1754*, *1755*, *2033*
<mood>, *1759*, *1900*, *2325*
@motivation
<annotation>, *632*, *971*, *1640*
<move>, *262*, *1762*
<msContents>, *361*, *376–381*, *400*, *401*, *531*, *1397*, *1764*, *1767*, *1770*, *1781*, *1826*
@msd
<pc>, *654*, *884*, *1897*
<w>, *654*, *884*, *1897*
<msDesc>, *360*, *361*, *400*, *401*, *428*, *1397*, *1767*, *1770*, *1781*
<msFrag>, *401*, *1770*
<msIdentifier>, *360*, *361*, *373–375*, *400*, *401*, *428*, *1137*, *1397*, *1555*, *1767*, *1770*, *1772*, *1781*, *2043*
<msItem>, *361*, *367–369*, *377–382*, *1397*, *1693*, *1696*, *1764*, *1767*, *1774*
<msItemStruct>, *1776*
<msName>, *373*, *374*, *401*, *1137*, *1770*, *1779*, *1781*
<msPart>, *400*, *1781*
<musicNotation>, *392*, *1784*, *1785*
@mutual
<relation>, *512*, *513*, *1667*, *1677*, *1886*, *2036*
@my:topic
<p>, *784*

@n
<ab>, *601*
<addSpan>, *431*, *941*
<app>, *467*
<author>, *836*
<body>, *134*
<cb>, *1082*
<change>, *67*, *1092*

<div1>, *132, 133, 136, 137, 170, 171, 177–179, 184, 191, 255, 256, 261, 267, 268, 575, 601, 602, 1258, 1262, 1509*
<div2>, *132, 133, 169–171, 177–179, 256, 267, 268, 601, 602, 1258, 1262, 1265*
<div3>, *132, 133, 1262, 1265, 1269*
<div4>, *1262, 1269*
, *6, 111, 134, 177, 179, 181, 184, 186, 187, 195, 234, 238, 256, 293, 304, 309, 310, 413, 476, 478, 554, 587, 631, 632, 908, 1254, 1640*
<divGen>, *129, 130, 1284*
<eTree>, *710, 711, 713, 715, 1350, 1418*
<edition>, *31, 1300, 1301*
<entry>, *321, 324, 328, 335, 343, 344, 2237*
<fLib>, *662, 665, 1402*
<figure>, *554, 556*
<forest>, *715, 1418*
<formula>, *552*
<front>, *134*
<fvLib>, *663, 665, 675, 676, 848, 1445*
<gb>, *1456*
<head>, *134*
<hom>, *324*
<item>, *6, 120, 199, 200, 587, 1006, 1566, 1635*
<l>, *6, 132, 133, 239–241, 437, 460, 466, 476–481, 483, 833, 2119*
<lb>, *437, 592, 721, 1510*
<lg>, *168, 169, 191, 234, 238, 241*
<metSym>, *244*
<milestone>, *50, 136–138, 889, 1735*
<msItem>, *367–369, 377–379, 381, 1693, 1764*
<msItemStruct>, *1776*
<note>, *124, 1809*
<p>, *134, 631, 632, 784, 1640*
<pb>, *196, 442, 857, 1456, 1894*
<ref>, *308, 613*
<region>, *2031*
<s>, *674, 676, 1629*
<said>, *724*
<seal>, *394, 2111*
<seg>, *240, 602, 631, 632, 725, 1640*
<sense>, *324, 326–328, 333–335, 339, 341, 344, 345, 349, 350, 352, 356, 1336, 2006–2008, 2127, 2237*
<spGrp>, *259, 2196*
<stage>, *2203*
<surface>, *454*
<surfaceGrp>, *406, 454, 1495, 1496, 2171*
<text>, *134*
<textDesc>, *569, 1076, 1965, 2299*
<titlePage>, *134*
<tree>, *705–708, 2338*
<name>, *8, 9, 12, 29–31, 35, 62, 65, 68, 81, 97, 100, 104–107, 109, 147, 148, 153, 154, 168, 171, 186, 202, 204, 249, 251, 252, 270, 276, 277, 302, 307, 308, 317, 318, 361, 369, 370, 380, 394, 396, 398, 424, 426, 440, 485, 486, 490, 495–497, 499, 510–512, 515, 524, 525, 538, 539, 547, 570, 571, 588, 589, 607, 619, 623, 687, 731, 835, 836, 838, 876, 897, 918, 935, 947, 1011, 1034, 1043, 1092, 1131, 1146, 1207, 1211, 1294, 1301, 1309, 1343, 1477, 1693, 1788, 1808, 1846, 1848, 1853, 1867, 1902, 1943, 1965, 2010, 2012, 2014, 2049, 2051, 2054, 2111, 2130, 2132, 2156, 2219, 2285, 2318, 2323, 2402*
@name
<attRef>, *772, 997*
<dataFacet>, *757, 1193*
<dataRef>, *757, 1193, 1195, 1198, 1200, 1235*
<equiv>, *746, 747, 759, 894, 1348, 2388*
<f>, *314, 541, 658–675, 683, 686, 687, 847, 848, 1022, 1024, 1144, 1219, 1376, 1377, 1402, 1430, 1538, 1540, 2208, 2223, 2261, 2305, 2393, 2397, 2404, 2407*
<fDecl>, *680, 682–684, 687, 850, 1386, 1388, 1434, 1436, 1438, 2399, 2411*
<localProp>, *209, 213, 216, 1096, 1483, 1687, 2361*
<namespace>, *849, 1792, 2269, 2271*
<param>, *764–766, 1742, 1745, 1748, 1881*
<relation>, *512, 513, 523, 524, 1362, 1667, 1677, 1886, 2036*
<unicodeProp>, *210, 216, 863, 1096, 1098, 1687, 2361*
<unihanProp>, *2369*
<vLabel>, *666, 2404*
<nameLink>, *493, 1791*
<namespace>, *849, 1792, 2269, 2271*
<nationality>, *510, 1795*
@new
<handShift>, *439, 460*
<shift>, *280, 284, 2148*
@next
<cl>, *638*
<def>, *356*
<lg>, *267*
<oRef>, *347*
<orth>, *355*
<pron>, *355*
<q>, *619*
<said>, *237*
<seg>, *725*
<node>, *691–697, 699, 702, 1492, 1797–1800*
@norm
<orth>, *354*
<pos>, *882*
<tns>, *354*
<usg>, *354*
<w>, *315, 316, 655, 883*
<normalization>, *43, 68, 575, 576, 1306, 1802, 2285*
@notAfter
<acquisition>, *396*

<affiliation>, *947*
<age>, *950*
<application>, *56*, *984*
<birth>, *488*
<change>, *68*
<custEvent>, *399*, *1185*
<date>, *114*, *459*, *487*, *489*, *737*, *1954*
<death>, *488*, *1910*
<education>, *1310*
<event>, *529*
<floruit>, *510*, *1408*
<orgName>, *514*, *2214*
<origDate>, *361*, *396*, *1781*, *1864*
<origin>, *396*, *1867*
<persName>, *512*
<placeName>, *517–520*
<population>, *523*
<residence>, *738*, *1886*, *2046*
<state>, *521*
<notatedMusic>, *553*, *554*, *1805*
@notation
<formula>, *550*, *1424*
<pRef>, *1353*
<pron>, *326*, *331–333*, *344*, *345*, *348*, *349*, *356–358*, *1972*
@notBefore
<acquisition>, *396*
<affiliation>, *947*
<birth>, *370*, *488*
<change>, *68*, *459*, *1092*
<custEvent>, *399*, *1185*
<date>, *114*, *487*, *737*, *1954*
<death>, *488*, *1910*
<education>, *1310*
<event>, *529*
<floruit>, *510*, *1408*
<licence>, *1624*
<nationality>, *510*, *1795*
<orgName>, *514*, *2214*
<origDate>, *361*, *396*, *1781*, *1864*
<origin>, *396*, *1867*
<persName>, *512*
<placeName>, *517–520*
<population>, *523*
<region>, *520*
<residence>, *489*, *2046*
<state>, *510*, *514*, *521*, *2214*
<note>, *27*, *36*, *67*, *68*, *116*, *118*, *123–125*, *145*, *146*, *150*, *153*, *154*, *157*, *158*, *163*, *250*, *251*, *277*, *293*, *309*, *311*, *344*, *377*, *399*, *424*, *426*, *427*, *466*, *482*, *483*, *502*, *503*, *507*, *580*, *582–584*, *731*, *832*, *871*, *892*, *971*, *980*, *1127*, *1362*, *1640*, *1696*, *1764*, *1808*, *1809*, *1811*, *1813*, *1910*, *1920*, *2012*, *2119*
<noteGrp>, *125*, *832*, *1811*
<notesStmt>, *27*, *36*, *68*, *1813*
@ns
<attDef>, *780*, *781*, *784*
<elementSpec>, *747*, *781–783*
<schemaSpec>, *785*
<num>, *100*, *110*, *113*, *517*, *547*, *844*, *935*, *1690*, *1711*, *1817*, *2374*, *2375*
<number>, *333*, *348–350*, *352*, *356*, *1063*, *1759*, *1819*, *1900*, *2325*
<numeric>, *314*, *541*, *660*, *669–671*, *1821*
<nym>, *536–539*, *1658*, *1823*
@nymRef
<forename>, *537*
<geogFeat>, *539*
<name>, *538*

<object>, *530*, *531*, *533*, *534*, *1661*, *1825*, *1826*, *1832*
<objectDesc>, *361*, *384*, *385*, *400*, *534*, *1767*, *1781*, *1830*, *1925*, *2243*
<objectIdentifier>, *530*, *531*, *533*, *534*, *1661*, *1825*, *1826*, *1831*, *1832*
<objectName>, *501*, *530*, *531*, *533*, *534*, *1661*, *1825*, *1826*, *1831*, *1832*, *1836*
<objectType>, *364*, *534*, *1838*
<occupation>, *370*, *537*, *570*, *1842*, *1886*
@occurs
<tagUsage>, *47*, *1792*, *2269*, *2271*
<offset>, *499*, *500*, *517*, *518*, *522*, *540*, *1690*, *1846*, *1932*
<opener>, *186*, *188*, *508*, *846*, *1848*, *1950*
@opt
<def>, *356*
<orth>, *355*
<pron>, *355*
@ord
<iNode>, *706*, *2338*
<root>, *706*, *2338*
<tree>, *705–708*, *1655*, *2338*
@order
<graph>, *691–697*, *699*, *702*, *1492*
<tree>, *705–708*, *2338*
@ordered
<listChange>, *458–460*, *1650*
<oRef>, *335*, *339*, *346*, *347*, *1063*, *1353*, *1759*, *1851*, *1900*, *2325*
<org>, *514*, *515*, *1664*, *1853*, *2214*
@org
<attList>, *756*, *995*
<div1>, *182*
, *274*
<vColl>, *667–669*, *671–673*, *2397*, *2405*
<vMerge>, *673*, *2405*
<orgName>, *36*, *156*, *158*, *495–497*, *501*, *508*, *514*, *515*, *524*, *533*, *631*, *947*, *1000*, *1664*, *1825*, *1836*, *1853*, *1858*, *2012*, *2214*
<orig>, *98*, *99*, *106*, *279*, *316*, *592*, *980*, *1100*, *1860*, *2027*, *2414*

@orig
<orth>, *354, 355, 2007, 2008*
<pron>, *355*
<tns>, *354*
<usg>, *354, 2007, 2008*
<w>, *655, 882, 883*
<origDate>, *361, 396, 531, 533, 534, 1661, 1781, 1825, 1826, 1864*
<origin>, *361, 396, 531, 533, 534, 1524, 1661, 1825, 1826, 1867*
@origin
<timeline>, *288–291, 605*
<origPlace>, *361, 396, 531, 533, 1661, 1781, 1825, 1826, 1870*
<orth>, *320–322, 326–328, 330–335, 337, 341–345, 347–358, 537, 849, 1063, 1103, 1140, 1217, 1237, 1336, 1339, 1353, 1421, 1459, 1486, 1489, 1526, 1529, 1569, 1593, 1613, 1759, 1819, 1823, 1851, 1873, 1900, 1939, 1972, 2006–2008, 2221, 2227, 2237, 2260, 2325, 2383, 2445*
@otherLangs
<textLang>, *163, 164, 382, 383, 2302*
@outDegree
<iNode>, *1553*
<node>, *693, 694, 696, 697, 699, 1799, 1800*
<root>, *2075*
@output
<model>, *765, 1742*
<modelGrp>, *763, 765, 1745, 1748*
<modelSequence>, *766*
<outputRendition>, *761–763, 1742, 1745, 1876*

<p>, *8–12, 22, 27, 33, 36, 40, 43, 49, 61, 63, 68, 74, 75, 86, 90–94, 98, 99, 101, 102, 104–107, 111–114, 116, 119, 125, 127, 128, 131, 134, 136, 137, 139, 141, 143, 169, 172, 178–182, 184–189, 191, 195, 196, 199, 200, 204, 213, 237, 244, 247–249, 251, 252, 256, 257, 261, 262, 265–271, 276, 277, 293, 299–302, 304, 306, 308–312, 360, 361, 364, 368, 372, 376, 382, 384–399, 413, 427, 442, 500, 501, 507, 508, 510, 512, 515, 521, 530, 531, 533, 534, 543, 550, 552, 554, 555, 570, 571, 575, 576, 587, 589, 590, 594, 595, 602, 608, 609, 612, 614, 618–620, 627, 631, 632, 636, 638, 643, 645–648, 687, 720, 721, 726, 734, 741–744, 747, 758, 784, 832, 838, 843, 845, 857, 865, 870, 871, 878, 884, 885, 897, 913, 927, 929, 943, 988, 991, 1004, 1006, 1014, 1029, 1030, 1052, 1053, 1055, 1060, 1100, 1123, 1131, 1135, 1146, 1167, 1191, 1207, 1212, 1215, 1227, 1254, 1258, 1262, 1265, 1269, 1272, 1277, 1281, 1284, 1302, 1306, 1312, 1334, 1346, 1362, 1367, 1394, 1397, 1405, 1426, 1479, 1501, 1509, 1510, 1519, 1524, 1531, 1564, 1572, 1587, 1595, 1605, 1607, 1624, 1631, 1636, 1640, 1661, 1672, 1678, 1708, 1712, 1720, 1762, 1764, 1767, 1784, 1802, 1811, 1817, 1825, 1826, 1830, 1832, 1836, 1838, 1848, 1878, 1886, 1894, 1902, 1903, 1909, 1925, 1943–1945, 1950, 1959, 1967, 1979, 1986, 1993, 1994, 1999, 2010, 2012, 2014, 2015, 2023, 2038, 2051, 2052, 2073, 2083, 2087, 2092–2094, 2098, 2105, 2111, 2113, 2124, 2135, 2139, 2165, 2169, 2173, 2180, 2185, 2192, 2196, 2208, 2210, 2217, 2280, 2285, 2318, 2342, 2349, 2375, 2388, 2402, 2418*
<param>, *764–766, 1742, 1745, 1748, 1881*
<paramList>, *766, 767, 1883, 1884*
<paramSpec>, *766, 767, 1883, 1884*
@parent
<iNode>, *705–708, 1553, 2338*
<leaf>, *705–708, 1614, 2338*
@part
<cl>, *637, 638*
, *181*
<l>, *101, 168–171, 232, 249, 250, 263, 264, 1341*
<lg>, *169, 265–268*
<seg>, *235, 600, 726*
@partial
<tagsDecl>, *84, 849, 2269, 2271*
<particDesc>, *309, 311, 570, 1886*
@parts
<nym>, *538*
@passive
<interaction>, *569, 1557, 2299*
<relation>, *512, 513, 523, 524, 1362, 1677, 2036*
<path>, *410, 411, 1889*
@pattern
<metDecl>, *241, 244, 245, 1730*
<pause>, *279, 281, 284, 286, 293, 294, 306, 1892, 2148*
<pb>, *118, 126, 196, 200, 204, 369, 404, 413, 442, 857, 1082, 1456, 1605, 1693, 1894*
<pc>, *77, 322, 642, 643, 654, 884, 885, 1897, 2208, 2414*
<per>, *1759, 1900, 2325*
@perf
<move>, *1762*
<performance>, *248, 251, 252, 1762, 1902, 1903*
@period
<placeName>, *488*
<persName>, *13, 36, 63–65, 67, 85, 144, 145, 148, 150, 151, 154, 159, 165, 309, 311, 370, 490–494, 503, 507–512, 514, 517, 525, 534, 537, 631, 632, 731, 733, 736, 739, 871, 897, 932, 971, 1000, 1088, 1169, 1171, 1172, 1356, 1362, 1416, 1465, 1640, 1690, 1720, 1755, 1791, 1853, 1867, 1886, 1907, 1909, 1910, 1913, 1920, 1943, 2052, 2054, 2073, 2214*
<person>, *85, 279, 282, 287, 293, 309, 311, 370, 430, 503, 507, 509–512, 525, 537, 570, 871, 1356, 1357, 1362, 1598, 1667, 1677, 1886, 1909, 1910, 1913, 1920, 1943, 2148*
<persona>, *1913*
<personGrp>, *1667, 1916*

<persPronouns>, *507*, *508*, *1920*
<phr>, *599*, *638*, *640*, *642*, *650*, *674*, *723*, *1048*, *1897*, *1923*, *1943*
<physDesc>, *361*, *384*, *400*, *531*, *533*, *534*, *1661*, *1708*, *1767*, *1781*, *1825*, *1826*, *1838*, *1925*
<place>, *502*, *517–525*, *631*, *632*, *871*, *1127*, *1235*, *1356*, *1361*, *1362*, *1670*, *1677*, *1689*, *1690*, *1927*, *2210*, *2293*
@place
<add>, *102*, *103*, *171*, *430*, *431*, *440*, *445*, *453*, *892*, *925*, *2230*
<fw>, *457*, *1448*
<head>, *1510*
<label>, *124*, *1587*
<metamark>, *452*
<note>, *116*, *123*, *124*, *153*, *154*, *157*, *158*, *250*, *251*, *580*, *582–584*, *892*, *1640*, *1808*
<rt>, *90–93*, *1999*, *2083*, *2087*
<stage>, *2203*
@placement
<punctuation>, *1986*
<placeName>, *13*, *54*, *64*, *65*, *101*, *188*, *488*, *491*, *497–503*, *510*, *517–525*, *533*, *631*, *632*, *731–733*, *846*, *859*, *871*, *1127*, *1235*, *1252*, *1356*, *1361*, *1362*, *1670*, *1689*, *1690*, *1711*, *1825*, *1836*, *1846*, *1858*, *1907*, *1910*, *1931*, *1932*, *1950*, *2012*, *2031*, *2046*, *2137*, *2142*, *2210*, *2293*, *2374*, *2379*
@points
<path>, *410*, *411*, *1889*
<zone>, *410*, *418*, *2448*, *2473*
<population>, *523*, *1935*, *2210*
<pos>, *326–328*, *330*, *332–334*, *345*, *349*, *350*, *352*, *353*, *356*, *357*, *848*, *849*, *881*, *882*, *1063*, *1217*, *1336*, *1459*, *1489*, *1526*, *1819*, *1939*, *1972*, *2006*, *2237*
@pos
<pc>, *654*, *884*, *885*, *1897*
<w>, *313*, *314*, *316–318*, *654*, *655*, *882–885*, *1897*, *2414*
<post>, *299–306*, *309–318*, *838*, *1943–1945*
<postBox>, *1947*
<postCode>, *109*, *306*, *935*, *1362*, *1949*
<postscript>, *188*, *1950*
@pre
<pc>, *643*
<precision>, *737*, *738*, *1954*
@precision
<precision>, *737*, *738*, *1954*
@predicate
<equiv>, *894*
<model>, *762–765*, *768*, *894*, *1741*, *1742*, *1745*, *1876*, *1881*
<pRef>, *348*, *1353*
@prefix
<schemaSpec>, *55*, *773*, *798*, *2103*
<prefixDef>, *589*, *590*, *849*, *1672*, *1959*
<preparedness>, *311*, *569*, *1076*, *1961*, *1965*, *2299*
@prev
<cl>, *638*
<def>, *356*
<lg>, *267*
<oRef>, *347*
<orth>, *355*
<pron>, *355*
<said>, *237*
<seg>, *725*
<principal>, *30*, *1963*
<profileDesc>, *61*, *62*, *68*, *309*, *311*, *458–460*, *913*, *1092*, *1650*, *1965*
<projectDesc>, *40*, *1967*, *2285*
<prologue>, *249*, *1969*
<pron>, *326–328*, *330–334*, *339*, *344*, *345*, *348–352*, *355–358*, *1217*, *1237*, *1336*, *1339*, *1526*, *1529*, *1569*, *1819*, *1972*, *2006*, *2237*
@property
<citeData>, *141*, *597*, *1105*
<provenance>, *361*, *396*, *531*, *533*, *1524*, *1661*, *1825*, *1826*, *1975*
<ptr>, *8*, *12*, *56*, *116–118*, *164*, *165*, *199*, *204*, *302*, *306–308*, *314*, *338*, *342*, *344*, *352*, *452*, *471*, *525*, *552–554*, *557*, *580*, *582*, *586*, *620*, *632*, *718*, *742*, *748*, *893*, *969*, *971*, *982*, *984*, *1107*, *1127*, *1170*, *1675*, *1680*, *1683*, *1805*, *1920*, *1977*, *1979*, *2335*, *2431*
<publicationStmt>, *22*, *27*, *33*, *34*, *68*, *144*, *306*, *307*, *687*, *1014*, *1394*, *1979*, *1984*, *2280*, *2285*
<publisher>, *10*, *33*, *34*, *38*, *46*, *68*, *144–146*, *148*, *150*, *153–155*, *157–159*, *161*, *162*, *165*, *202*, *307*, *308*, *857*, *878*, *1011*, *1014*, *1019*, *1543*, *1647*, *1755*, *1979*, *1982*, *1984*, *2033*, *2109*, *2169*
<pubPlace>, *10*, *33*, *34*, *38*, *46*, *68*, *144–148*, *150*, *153–155*, *157–159*, *161*, *162*, *165*, *202*, *307*, *398*, *878*, *1011*, *1014*, *1019*, *1294*, *1543*, *1755*, *1979*, *1982*, *1984*, *2033*
<punctuation>, *1986*
<purpose>, *311*, *569*, *1076*, *1965*, *1988*, *2299*

<q>, *8*, *47*, *75*, *79–81*, *89*, *104*, *105*, *172*, *187*, *189*, *199*, *204*, *347*, *424*, *459*, *510*, *530*, *594*, *598*, *614*, *616*, *618–620*, *844*, *929*, *1233*, *1331*, *1426*, *1519*, *1572*, *1878*, *1991*, *2027*, *2080*, *2111*, *2388*
@quantity
<depth>, *1229*, *1714*
<dim>, *366*, *534*, *738*
<gap>, *99*, *100*, *294*, *317*, *318*, *381*, *436*, *443*, *592*, *1224*, *1453*, *2230*
<height>, *366*, *1243*, *1517*, *1714*
<measure>, *32*, *111*, *112*, *307*, *386*, *500*, *525*, *886*, *1375*, *1711*
<population>, *2210*
<space>, *455*, *2177*
<width>, *366*, *1243*, *1714*

@queryBinding
<constraintDecl>, *754*, *1153*
<quotation>, *43*, *68*, *83*, *1306*, *1993*, *1994*, *2093*, *2094*
<quote>, *xxv*, *10*, *11*, *86*, *87*, *187*, *195*, *196*, *202*, *322*, *334–337*, *339*, *341*, *342*, *344*, *347*, *361*, *372*, *393–395*, *471*, *511*, *531*, *580*, *582–584*, *878*, *1103*, *1341*, *1426*, *1586*, *1826*, *1851*, *1920*, *1945*, *1986*, *1997*, *2127*

<rb>, *90–94*, *1999*, *2083*, *2087*
<rdg>, *428*, *435*, *463–470*, *472*, *475–483*, *631*, *979*, *980*, *1589*, *1590*, *1617*, *1644*, *2001*, *2003*, *2119*, *2425*, *2428*, *2429*, *2433*
<rdgGrp>, *466–468*, *979*, *2003*
@real
<l>, *239*, *1583*
<seg>, *240*
@reason
<gap>, *99*, *100*, *181*, *294*, *317*, *318*, *381*, *436*, *443*, *592*, *838*, *895*, *898*, *1224*, *1453*, *1774*, *1944*, *2230*
<supplied>, *317*, *318*, *438*, *443*, *455*, *592*, *1776*, *2051*, *2241*
<surplus>, *437*, *2255*
<unclear>, *102*, *294*, *443*, *2349*, *2351*
<recordHist>, *397*, *398*, *927*, *943*, *2010*
<recording>, *276*, *277*, *2012*, *2014*
<recordingStmt>, *276*, *2014*, *2015*
<redo>, *2017*
<ref>, *63*, *86*, *109*, *117–119*, *125*, *126*, *149*, *150*, *161*, *164*, *165*, *200*, *203*, *204*, *301*, *302*, *308–311*, *314*, *335*, *343*, *344*, *349*, *350*, *381*, *385*, *389*, *394*, *398*, *459*, *470–472*, *487*, *524*, *534*, *556*, *582*, *583*, *587*, *595*, *613*, *865*, *880*, *938*, *1171*, *1172*, *1262*, *1353*, *1362*, *1397*, *1586*, *1593*, *1652*, *1776*, *1910*, *1920*, *1944*, *1945*, *1997*, *2010*, *2021*, *2111*, *2167*, *2445*
@ref
<actor>, *253*, *922*
<author>, *151*, *153*
<bibl>, *143*, *165*
<editor>, *153*
<event>, *524*
<eventName>, *500*, *1362*
<faith>, *1384*
<forename>, *537*
<g>, *128*, *213*, *215*, *216*, *293*, *379*, *421–423*, *467*, *468*, *482*, *601*, *642*, *964*, *979*, *1450*, *1536*, *1590*, *2429*
<geogName>, *498*, *499*, *539*
<name>, *106*, *302*, *317*, *318*, *370*, *486*, *490*, *495*, *497*, *499*, *510*, *511*, *588*, *589*, *835*, *836*, *1788*
<objectName>, *501*, *533*, *1825*, *1836*
<orgName>, *495*, *501*, *533*, *631*, *1825*, *1836*
<persName>, *67*, *490*, *492–494*, *525*, *631*, *2052*, *2073*
<placeName>, *54*, *488*, *497–501*, *533*, *631*, *859*, *871*, *1711*, *1825*, *1836*, *2210*, *2374*, *2379*
<principal>, *1963*
<relation>, *1362*, *2036*
<resp>, *2049*
<rs>, *106*, *490*, *495*, *497*, *500*, *1362*
<socecStatus>, *506*
<sponsor>, *2199*
<state>, *2214*
<surname>, *491*
<term>, *88*, *2289*, *2290*
<refsDecl>, *49*, *50*, *139–141*, *594–596*, *2022*, *2023*, *2285*
<refState>, *50*, *141*, *2025*
<reg>, *98*, *99*, *106*, *279*, *316*, *592*, *980*, *1100*, *1860*, *2027*, *2414*
<region>, *373–375*, *498*, *500*, *510*, *518*, *520*, *525*, *533*, *534*, *1137*, *1661*, *1825*, *1831*, *1931*, *1932*, *2031*, *2046*, *2142*, *2210*
<relatedItem>, *161–163*, *308*, *2033*
<relation>, *512*, *513*, *523*, *524*, *1362*, *1667*, *1677*, *1886*, *2036*
<remarks>, *747*, *758*, *1123*, *2038*, *2388*
@rend
<castGroup>, *253*, *254*, *1065*
<date>, *615*
<del>, *103*, *430*, *451*, *453*, *456*, *895*, *1224*, *2354*
<delSpan>, *432*, *433*
<emph>, *8*, *80*, *645*
<figure>, *587*
<foreign>, *89*
<formula>, *552*
<gloss>, *88*, *89*
<graph>, *1492*
<head>, *171*, *184*, *189*, *254*, *546*, *602*, *873*, *1070*
<headItem>, *1512*, *1514*
<headLabel>, *1512*, *1514*
<hi>, *81*, *89*, *344*, *424*, *456*, *470*, *472*, *601*, *873*, *1082*, *1424*, *1510*, *1522*, *1817*
<l>, *101*, *168*
<lg>, *1644*
<list>, *6*, *119*, *120*, *122*, *587*, *1314*, *1510*, *1566*, *1586*, *1634–1636*
<metamark>, *101*, *448*, *452*, *1728*
<mod>, *446*, *2017*
<name>, *8*, *251*
<note>, *251*
<p>, *413*, *2052*
<persName>, *739*
<post>, *302*, *310*, *315–318*, *838*, *1943*, *1944*
<pron>, *351*
<ptr>, *582*
<q>, *89*, *614*
<ref>, *582*, *583*
<rubric>, *369*, *1693*
<said>, *84*, *2094*
<seg>, *2122*
<signed>, *301*, *309*, *310*, *838*, *1944*
<soCalled>, *645*
<spGrp>, *269*, *2052*

<speaker>, *2052*
<stage>, *170, 171, 602, 2052, 2202*
<table>, *546*
<term>, *88, 118, 1482, 2289, 2290*
<title>, *89*
<undo>, *451, 2354*
<val>, *1314*
<rendition>, *9, 44–46, 81, 83, 84, 874, 2040, 2041, 2224, 2269, 2271*
@rendition
<docDate>, *46*
<docImprint>, *46*
<docTitle>, *46*
<emph>, *81*
<head>, *874*
<hi>, *46, 874*
<name>, *81*
<p>, *9*
<pubPlace>, *46*
<publisher>, *46*
<q>, *47*
<said>, *83*
<soCalled>, *2210*
<titlePart>, *46*
@replacementPattern
<cRefPattern>, *139, 140, 594, 595, 1180, 2022, 2285*
<prefixDef>, *589, 590, 849, 1672, 1959*
@replyTo
<post>, *299, 309, 310, 1945*
<repository>, *360, 361, 373–375, 400, 401, 428, 533, 961, 1137, 1555, 1661, 1767, 1770, 1772, 1781, 1825, 1831, 2043*
@require
<anyElement>, *976, 977*
<residence>, *488, 489, 510, 570, 738, 1886, 1920, 2046*
<resp>, *12, 29–31, 35, 36, 67, 68, 97, 148, 153, 154, 276, 277, 307, 380, 424, 440, 632, 687, 876, 971, 1043, 1092, 1301, 1640, 1808, 2012, 2014, 2049, 2051, 2052, 2054, 2130, 2132, 2285, 2323*
@resp
<abstract>, *913*
<add>, *440*
<corr>, *11, 12, 97, 426, 428, 440, 441, 876*
<delSpan>, *432*
<desc>, *1324*
<ellipsis>, *1324*
<event>, *512*
<ex>, *424*
<expan>, *424*
<gap>, *99*
<handShift>, *439*
<interpGrp>, *648, 1562*
<lem>, *467, 468, 979*
<note>, *124, 731, 1640, 1808*
<origin>, *1867*
<persName>, *736, 1910*
<population>, *523, 1935*
<rdg>, *980*
<reg>, *2027*
<relation>, *2036*
<respons>, *441, 739, 2051, 2052*
<spanGrp>, *646*
<supplied>, *438, 443, 455*
<witDetail>, *2428*
<respons>, *441, 739, 2051, 2052*
<respStmt>, *12, 29–31, 35, 36, 67, 68, 97, 148, 153, 154, 276, 277, 307, 380, 424, 440, 632, 687, 876, 971, 1043, 1092, 1301, 1640, 1808, 2012, 2014, 2049, 2051, 2052, 2054, 2130, 2132, 2285, 2323*
<restore>, *435, 2058*
@restriction
<dataRef>, *1195, 1235*
@result
<join>, *237, 617, 618, 726, 1572*
<joinGrp>, *618, 620, 1575*
<retrace>, *448–450, 460, 1728*
<revisionDesc>, *67, 68, 572, 632, 855, 971, 1092, 1650, 2063, 2285*
<rhyme>, *242, 243, 2066*
@rhyme
, *238, 241*
<lg>, *239, 241–243, 1640, 2066*
<role>, *251–254, 257, 258, 262, 833, 922, 1065, 1068, 1070, 1762, 1902, 2068*
@role
<cell>, *546–548, 1085, 1127, 2077, 2264*
<editor>, *162, 1305*
<name>, *485*
<persName>, *534*
<person>, *309, 311, 1598, 1909, 1910*
<personGrp>, *1916*
<roleName>, *2073*
<row>, *547, 2077, 2264*
<roleDesc>, *253, 254, 833, 922, 1065, 1070, 2068, 2070*
<roleName>, *13, 493, 494, 508, 932, 1416, 1907, 1920, 2073*
<root>, *705–708, 1655, 2075, 2338*
<row>, *100, 546–549, 1085, 1127, 2077, 2264*
@rows
<table>, *100, 546, 547, 2264*
<rs>, *58, 104–106, 251, 490, 495, 497, 500, 1362, 1629, 1902, 1903, 2080*
<rt>, *90–94, 1999, 2083, 2087*
<rubric>, *361, 364, 369, 1373, 1693, 1776, 2085, 2206*
<ruby>, *90–94, 1999, 2083, 2087*
@ruledLines
<layout>, *388, 1605, 1607, 1830*

<s>, *218*, *225*, *491*, *600*, *610*, *616*, *623*, *636*, *638*, *641*, *644–651*, *654*, *674*, *676*, *884*, *885*, *969*, *1572*, *1629*, *1631*, *1897*, *2091*, *2122*
<said>, *83–86*, *237*, *511*, *721*, *724*, *2092–2094*
<salute>, *186*, *188*, *204*, *508*, *846*, *1006*, *1131*, *1848*, *1950*, *2097*
@sameAs
<date>, *615*
<name>, *65*
<persName>, *65*
@sample
, *181*, *274*
<samplingDecl>, *40*, *68*, *308*, *309*, *2098*
@scale
<graphic>, *1393*
<schemaRef>, *56*, *2100*
<schemaSpec>, *3*, *4*, *55*, *260*, *770–773*, *778*, *785*, *798*, *878*, *1120*, *1195*, *1700*, *2103*
@scheme
<att>, *991*
<catRef>, *60*, *1080*
<classCode>, *60*, *68*, *156*, *158*, *1115*, *2298*
<constraintDecl>, *754*, *1153*
<constraintSpec>, *754–756*, *759*, *1150*, *1156*
<gi>, *1479*
<keywords>, *59*, *68*, *1116*, *1577*, *2298*
<locus>, *369*, *387*
<locusGrp>, *369*
<occupation>, *1842*
<rendition>, *9*, *44*, *46*, *81*, *84*, *874*, *2040*, *2041*, *2269*, *2271*
<socecStatus>, *570*, *2162*
<styleDefDecl>, *45*, *2224*
@schemeVersion
<styleDefDecl>, *45*, *2224*
@scope
<dimensions>, *361*, *1925*
<handNote>, *389*, *897*, *1500*, *1503*
<height>, *366*, *1243*, *1714*
<join>, *237*, *618*, *726*, *1572*
<language>, *897*
<outputRendition>, *1745*
<rendition>, *46*, *83*, *84*, *2041*
<scriptNote>, *2108*
<typeNote>, *2344*
<width>, *366*, *737*, *738*, *1243*, *1714*
@scribe
<handNote>, *431*, *941*, *1503*
@scribeRef
<handNote>, *1503*
@script
<handNote>, *439*, *1505*
<scriptDesc>, *2105*
<scriptNote>, *2105*, *2108*
<scriptStmt>, *275*, *2109*
<seal>, *394*, *2111*, *2113*
<sealDesc>, *394*, *2113*
<secFol>, *371*, *2116*
<secl>, *437*, *2119*
<seg>, *227*, *234*, *235*, *237*, *240*, *279*, *286*, *287*, *451–453*, *481*, *537*, *598–600*, *602*, *604*, *607*, *609*, *612*, *613*, *615*, *622*, *625*, *626*, *631*, *632*, *641*, *647*, *648*, *720–722*, *725*, *726*, *974*, *1640*, *1945*, *2122*, *2210*, *2354*
<segmentation>, *43*, *2124*
@select
, *622*, *623*
<seg>, *622*
<u>, *622*
@selector
<rendition>, *9*, *44*, *84*, *2040*, *2269*, *2271*
<sense>, *322*, *324*, *326–328*, *333–335*, *337–339*, *341*, *344*, *345*, *349*, *350*, *352*, *356*, *1082*, *1336*, *2006–2008*, *2127*, *2237*
@seq
<add>, *433*
<del>, *433*
<sequence>, *751*, *752*, *769*, *795–797*, *1159*, *2129*
<series>, *147*, *148*, *158*, *159*, *162*, *277*, *307*, *1755*, *2012*, *2033*, *2130*
<seriesStmt>, *27*, *35*, *2132*
<set>, *247–249*, *2135*
<setting>, *571*, *1965*, *2137*
<settingDesc>, *571*, *1965*, *2139*
<settlement>, *109*, *360*, *361*, *373–375*, *400*, *401*, *428*, *498*, *500*, *517–520*, *525*, *531*, *533*, *534*, *935*, *961*, *1137*, *1169*, *1172*, *1252*, *1362*, *1397*, *1555*, *1661*, *1690*, *1767*, *1770*, *1772*, *1781*, *1825*, *1826*, *1831*, *1886*, *1910*, *1927*, *1931*, *1932*, *2043*, *2046*, *2142*
<sex>, *508*, *1920*, *2145*
@sex
<person>, *309*, *570*, *1598*, *1667*, *1886*, *1909*, *1910*, *1913*
<personGrp>, *1916*
<persona>, *1913*
<shift>, *280*, *284*, *2148*
<sic>, *11–13*, *97*, *294*, *393*, *424–428*, *436*, *437*, *440*, *441*, *876*, *1100*, *1165*, *2151*, *2210*
<signatures>, *371*, *2154*
<signed>, *185*, *186*, *188*, *199*, *204*, *301*, *309–312*, *838*, *1131*, *1944*, *1945*, *1950*, *2156*
@since
<when>, *288–290*, *605*, *2312*, *2420*
@size
<graph>, *691–697*, *699*, *702*, *1492*
<personGrp>, *1916*
<soCalled>, *87*, *182*, *283*, *471*, *645*, *991*, *995*, *2159*, *2210*
<socecStatus>, *506*, *570*, *2162*
@social
<distinct>, *82*

@sort
<forename>, *492*, *494*
<surname>, *492*, *494*
@sortKey
<term>, *127*, *128*, *899*
<sound>, *271*, *1060*, *2165*
<source>, *397*, *398*, *943*, *2010*, *2167*
@source
<corr>, *428*
<derivation>, *1232*
<egXML>, *1315*
<elementRef>, *11*, *878*, *1318*, *2103*
<lem>, *483*
<moduleRef>, *2103*
<normalization>, *43*, *1802*
<note>, *466*
<population>, *2210*
<quote>, *10*, *11*, *87*, *878*
<rdg>, *465*, *482*, *483*, *2119*
<schemaSpec>, *55*, *878*, *1195*, *1700*, *2103*
<supplied>, *438*, *443*, *455*, *2241*
<writing>, *283*
<sourceDesc>, *22*, *27*, *36*, *38*, *68*, *275*, *283*, *306–308*, *687*, *1014*, *1394*, *2169*, *2280*, *2285*
<sourceDoc>, *414*, *454*, *460*, *2171*, *2249*
<sp>, *169–172*, *249*, *250*, *256–259*, *261–271*, *437*, *602*, *615*, *618*, *833*, *1060*, *1343*, *1405*, *1572*, *1969*, *2052*, *2165*, *2173*, *2185*, *2196*, *2402*
<space>, *455*, *2177*
@space
<distinct>, *82*
, *644–646*, *655*, *1640*, *2180*, *2183*, *2208*
<spanGrp>, *646*, *655*, *1640*, *2183*, *2208*
@spanTo
<addSpan>, *431*, *447*, *941*
<damageSpan>, *442*, *443*, *1191*
<delSpan>, *432*, *433*, *447*, *1227*
<index>, *127*
<mod>, *446*, *2017*
<undo>, *451*
<speaker>, *169–172*, *250*, *256*, *257*, *259*, *261–271*, *602*, *615*, *618*, *833*, *1060*, *1572*, *2052*, *2165*, *2173*, *2185*, *2196*, *2402*
<specDesc>, *742*, *2186*, *2187*, *2194*
<specGrp>, *743*, *2189*, *2192*
<specGrpRef>, *744*, *2192*
<specList>, *742*, *2194*
<spGrp>, *259*, *264*, *265*, *269*, *2052*, *2196*
@split
<orth>, *355*
<pron>, *355*
<sponsor>, *2199*
<stage>, *168–172*, *249*, *250*, *256*, *258*, *260–262*, *265–269*, *602*, *833*, *1070*, *1343*, *1762*, *2052*, *2185*, *2196*, *2202*, *2203*
<stamp>, *364*, *2206*
<standOff>, *85*, *503*, *524*, *525*, *631*, *632*, *1640*, *2208*, *2210*
@start
<annotationBlock>, *974*
<kinesic>, *306*
<post>, *306*
<schemaSpec>, *3*, *4*, *773*, *1120*, *2103*
<surface>, *413*
<u>, *289*, *290*, *306*
<state>, *510*, *514*, *521*, *2214*
@status
<availability>, *34*, *1004*
<biblStruct>, *156*, *158*
<change>, *632*, *855*, *971*, *1092*
<correction>, *68*
<del>, *434*
<revisionDesc>, *855*, *1092*, *2063*
@stdDeviation
<precision>, *738*
<stdVals>, *68*, *2217*
@streams
<layout>, *1605*
<street>, *109*, *531*, *935*, *1826*, *1886*, *2219*
<stress>, *2221*
<string>, *541*, *659*, *660*, *664*, *665*, *667*, *671*, *684*, *687*, *2208*, *2223*, *2397*
@style
<ab>, *226*, *601*
<cell>, *224*
<emph>, *9*, *81*, *1920*
<head>, *874*
<hi>, *762*, *874*
<lg>, *219*, *227*
<list>, *221*
<name>, *9*
<p>, *91–94*, *2083*
<rt>, *94*
<s>, *225*
<seg>, *227*
<term>, *225*
<styleDefDecl>, *45*, *2224*
<subc>, *327*, *328*, *332*, *1489*, *1526*, *2227*, *2237*
<subst>, *431*, *433*, *434*, *445*, *460*, *2229*, *2230*
<substJoin>, *434*, *2233*
@subtype
<anchor>, *721*, *723*
<bibl>, *147*, *1011*
<correspAction>, *64*
, *274*, *632*, *1640*
<entry>, *345*
<graphic>, *1496*
<kinesic>, *306*
<ptr>, *1170*
<seal>, *394*, *2111*

<seg>, *599*, *600*, *974*
<w>, *1943*, *1944*
<summary>, *378*, *390*, *400*, *401*, *1770*, *1781*, *2105*, *2235*, *2342*
<superEntry>, *2237*
<supplied>, *100*, *101*, *317*, *318*, *425*, *438*, *443*, *455*, *592*, *1188*, *1325*, *1510*, *1776*, *2051*, *2241*
<support>, *361*, *364*, *385*, *1925*, *2243*, *2244*, *2418*
<supportDesc>, *361*, *385*, *387*, *1830*, *1925*, *2243*, *2244*
<surface>, *405–408*, *410*, *411*, *413–415*, *417–419*, *448*, *454*, *460*, *1378*, *1495*, *1496*, *1626*, *1728*, *1889*, *2171*, *2247*, *2249*, *2448*, *2473*
<surfaceGrp>, *406*, *454*, *1495*, *1496*, *2171*, *2249*
<surname>, *13*, *144*, *145*, *147–151*, *155*, *159*, *165*, *370*, *490–494*, *496*, *507*, *511*, *1011*, *1019*, *1362*, *1416*, *1465*, *1755*, *1791*, *1886*, *1907*, *2073*, *2252*
<surplus>, *437*, *2255*
<surrogates>, *399*, *927*, *2257*
<syll>, *330*, *2260*
<symbol>, *659*, *661*, *663*, *664*, *666–668*, *672–675*, *683*, *684*, *686*, *687*, *847*, *848*, *850*, *1022*, *1144*, *1376*, *1402*, *1430*, *1434*, *1438*, *1445*, *1538*, *1540*, *2208*, *2261*, *2305*, *2393*, *2397*, *2404*, *2405*, *2407*
@synch
<anchor>, *288–290*, *603*
<date>, *312*
<post>, *299*, *301*, *302*, *305*, *310*, *312–318*, *838*, *1944*, *1945*
<seg>, *604*
<u>, *604*
<unclear>, *604*
<when>, *289*, *291*

<table>, *100*, *546–549*, *1127*, *2264*
<tag>, *741*, *2267*
@tag
<langKnown>, *507*, *570*, *1596*, *1598*
@tags
<langKnowledge>, *507*, *570*, *1595*
<tagsDecl>, *9*, *45*, *84*, *849*, *2041*, *2224*, *2269*, *2271*
<tagUsage>, *47*, *260*, *849*, *1792*, *2269*, *2271*
@target
<alt>, *453*, *525*, *624–626*, *953*, *957*
<annotation>, *632*, *971*, *1640*
<calendar>, *1052*, *1055*
<catRef>, *52*, *60*, *1080*, *2298*
<certainty>, *441*, *456*, *731–733*, *735*, *736*, *897*, *1087*, *1088*
<change>, *460*
<citedRange>, *1107*
<fsdLink>, *679*, *1434*, *1440*
<gloss>, *88*, *1482*, *1725*, *2289*
<join>, *237*, *268*, *617*, *618*, *620*, *648*, *726*, *1572*, *1575*
<licence>, *34*, *632*, *971*, *1004*, *1623*, *1624*
<link>, *243*, *289*, *291*, *557*, *580*, *583–586*, *604*, *606–609*, *613*, *614*, *617*, *624*, *648*, *652*, *676*, *1629*, *1631*
<locus>, *369*, *1693*
<metamark>, *447*, *448*, *452*, *1728*
<note>, *424*, *426*, *427*, *482*, *483*, *582*, *731*, *1640*
<oRef>, *347*
<pRef>, *348*
<precision>, *737*, *738*, *1954*
<ptr>, *8*, *12*, *56*, *116–118*, *164*, *165*, *199*, *204*, *302*, *306–308*, *314*, *338*, *342*, *344*, *352*, *452*, *471*, *525*, *552–554*, *557*, *580*, *582*, *586*, *620*, *632*, *718*, *742*, *748*, *893*, *969*, *971*, *982*, *984*, *1107*, *1127*, *1170*, *1675*, *1680*, *1683*, *1805*, *1920*, *1977*, *1979*, *2335*, *2431*
<redo>, *2017*
<ref>, *63*, *109*, *117–119*, *126*, *149*, *150*, *161*, *164*, *165*, *200*, *204*, *301*, *302*, *308–311*, *314*, *335*, *343*, *344*, *349*, *398*, *459*, *470*, *472*, *487*, *534*, *556*, *582*, *583*, *587*, *613*, *865*, *880*, *938*, *1171*, *1172*, *1262*, *1362*, *1397*, *1652*, *1776*, *1910*, *1920*, *1944*, *1945*, *2010*, *2021*
<relatedItem>, *163*
<respons>, *441*, *739*, *2051*, *2052*
<rt>, *93*
, *645*, *2208*
<specGrpRef>, *744*, *2192*
<substJoin>, *434*, *2233*
<undo>, *451*, *2354*
<witDetail>, *470*, *2428*
@targetDatcat
<fDecl>, *684*, *850*
<tagUsage>, *849*
@targetEnd
<noteGrp>, *832*
@targetLang
<ptr>, *8*, *308*, *893*
<ref>, *308*
@targFunc
<linkGrp>, *585*, *604*, *606*, *607*, *648*
<taxonomy>, *50*, *51*, *59*, *68*, *382*, *487*, *849*, *851*, *1080*, *1116*, *2273*, *2274*, *2298*
<tech>, *2278*
<TEI>, *7*, *8*, *22*, *23*, *175*, *303*, *404*, *563*, *564*, *572*, *631*, *632*, *678*, *679*, *687*, *2208*, *2210*, *2280*, *2283*
<teiCorpus>, *23*, *303*, *563*, *564*, *2283*
<teiHeader>, *7*, *8*, *22*, *23*, *27*, *68*, *175*, *303*, *404*, *563*, *564*, *572*, *631*, *632*, *678*, *679*, *687*, *2208*, *2210*, *2280*, *2283*, *2285*
<term>, *8*, *59*, *68*, *80*, *88*, *107*, *118*, *121*, *127–129*, *178*, *218*, *225*, *389*, *391*, *849*, *851*, *899*, *1116*, *1269*, *1482*, *1500*, *1551*, *1577*, *1586*, *1785*, *1808*, *2021*, *2289*, *2290*
@terminal
<metSym>, *241*, *244*, *1733*
<terrain>, *521*, *1235*, *2293*
<text>, *7–9*, *22*, *23*, *134*, *136*, *137*, *175*, *189*, *191*, *193*, *194*, *231*, *233*, *249*, *268*, *303*, *306*, *404*, *413*, *460*, *476*,

*563*, *564*, *572*, *575*, *619*, *620*, *628*, *631*, *632*, *678*, *679*, *870*, *1315*, *1498*, *1640*, *2208*, *2210*, *2280*, *2283*, *2295*
<textClass>, *68*, *1116*, *2298*
<textDesc>, *311*, *569*, *1076*, *1965*, *2299*
<textLang>, *163*, *164*, *361*, *382*, *383*, *400*, *1774*, *1776*, *1781*, *2302*
<textNode>, *779*, *799*, *977*, *1321*, *1702*, *2304*
<then>, *683*, *686*, *687*, *1144*, *1432*, *1538*, *2305*
<time>, *114*, *301*, *302*, *310*, *314–316*, *540–542*, *571*, *838*, *844–846*, *1944*, *1945*, *1965*, *2309*, *2461*, *2462*
@time
<distinct>, *82*
<timeline>, *288–291*, *311*, *605*, *2312*
<title>, *10*, *22*, *27*, *29*, *30*, *35*, *36*, *38*, *50*, *59*, *67*, *68*, *86*, *87*, *89*, *114*, *124*, *143–155*, *157–163*, *165*, *172*, *189*, *199*, *200*, *252*, *275*, *277*, *306–308*, *361*, *367–369*, *377–381*, *389*, *397–399*, *426*, *459*, *503*, *511*, *519*, *531*, *534*, *587*, *607*, *632*, *687*, *758*, *857*, *865*, *878*, *966*, *1011*, *1014*, *1019*, *1043*, *1092*, *1146*, *1232*, *1394*, *1397*, *1647*, *1693*, *1754*, *1755*, *1764*, *1767*, *1774*, *1776*, *1781*, *1826*, *1853*, *1910*, *2012*, *2033*, *2109*, *2130*, *2132*, *2169*, *2257*, *2280*, *2285*, *2316*, *2323*
<titlePage>, *46*, *134*, *187*, *191*, *202*, *248*, *1287*, *2318*
<titlePart>, *46*, *134*, *187*, *189*, *191*, *202*, *248*, *266*, *1287*, *1296*, *1610*, *2295*, *2318*, *2321*
<titleStmt>, *22*, *27*, *29*, *30*, *36*, *67*, *68*, *124*, *144*, *306*, *307*, *632*, *687*, *1014*, *1092*, *1394*, *2280*, *2285*, *2323*
<tns>, *354*, *1759*, *1900*, *2325*
@to
<affiliation>, *947*
<app>, *478*, *479*
<arc>, *691*, *693*, *695–697*, *699*, *702*, *986*, *1492*
<biblScope>, *145*, *147*, *148*, *152*, *159*, *160*, *165*, *1017*, *1754*
<citedRange>, *1107*
<conversion>, *54*, *55*, *1162*
<custEvent>, *534*
<date>, *114*, *542*, *846*, *865*, *1172*
<education>, *1309*
<event>, *488*, *1356*, *1361*, *1362*, *1652*
<locus>, *367*, *368*, *379*, *381*, *388*, *389*, *1607*, *1693*, *1696*, *1776*
<persName>, *154*
<residence>, *488*, *489*
<rt>, *92*
, *644–646*, *655*, *1640*, *2180*, *2183*
<state>, *510*, *514*, *2214*
@to-custom
<event>, *843*
@toUnit
<conversion>, *54*, *55*, *859*, *860*, *1162*, *2377*, *2379*
@toWhom
<said>, *85*
<sp>, *833*
<stage>, *833*
<trailer>, *178*, *184*, *200*, *204*, *249*, *314*, *549*, *1040*, *1509*, *1969*, *2328*
<trait>, *509*, *2332*
@trans
<u>, *279*, *280*, *288*, *293*, *2347*
<transcriptionDesc>, *277*, *2334*
<transpose>, *452*, *1680*, *2335*
<tree>, *705–708*, *1655*, *2338*
<triangle>, *710*, *2340*
@trunc
<numeric>, *660*, *1821*
@type
<ab>, *525*, *556*, *620*
<abbr>, *115*, *472*
<addName>, *492*, *494*, *932*
<altGrp>, *525*
<altIdentifier>, *360*, *361*, *375*, *1781*
<anchor>, *721*, *723*
<app>, *467–469*, *979*, *2428*
<bibl>, *147*, *1011*
<biblStruct>, *156*, *158*, *165*, *1019*, *1755*
<bloc>, *498*, *1038*
<c>, *313*, *642*
<calendar>, *63*
<camera>, *1058*
<castItem>, *253*, *833*, *1068*
<cb>, *1605*
<cit>, *322*, *334–337*, *339*, *341*, *342*, *344*, *347*, *1103*, *1851*, *2127*
<cl>, *638*, *1113*
<classSpec>, *768*, *782*, *800*, *801*, *1123*
<constitution>, *569*, *1076*, *1149*, *1965*, *2299*
<corr>, *427*
<correspAction>, *64*, *65*, *1169*, *1172*
<country>, *498*
<custEvent>, *399*, *1183*, *1185*
<date>, *156*, *158*, *308*, *540*, *541*
<decoNote>, *391*, *1215*
<del>, *286*, *294*, *1224*
<derivation>, *569*, *1076*, *1231*, *1232*, *1965*, *2299*
<desc>, *301*, *313*, *1233*
<dim>, *366*, *534*, *738*, *1240*, *1243*
<dimensions>, *361*, *366*, *1243*, *1925*
<distinct>, *82*, *1246*
<district>, *498*, *517*, *935*, *1252*, *1690*
<div1>, *132*, *133*, *136*, *137*, *170*, *171*, *177–179*, *182*, *191*, *200*, *249*, *255*, *256*, *261*, *267*, *268*, *557*, *601*, *602*, *1258*, *1262*, *1284*, *1509*, *1635*
<div2>, *132*, *133*, *169–171*, *177–179*, *182*, *256*, *267*, *268*, *601*, *602*, *1258*, *1262*, *1265*, *1272*, *1277*, *1281*, *1509*
<div3>, *132*, *133*, *1262*, *1265*, *1269*
<div4>, *1262*, *1269*, *1272*

<div5>, *1272*, *1277*, *1281*
<div6>, *1281*
, *6*, *86*, *125*, *129*, *177*, *179*, *181*, *184–188*, *191*, *195*, *199*, *200*, *203*, *204*, *234*, *238*, *241*, *248*, *256*, *274*, *293*, *300*, *304*, *309*, *310*, *476*, *478*, *508*, *587*, *592*, *603*, *608–610*, *612*, *613*, *621–624*, *626*, *631*, *632*, *637*, *865*, *870*, *904*, *908*, *1006*, *1131*, *1207*, *1254*, *1405*, *1426*, *1510*, *1605*, *1631*, *1640*, *1950*, *2402*
<divGen>, *129*, *130*, *198*, *1284*
<domain>, *569*, *1076*, *1298*, *1965*, *2299*
<eLeaf>, *717*, *718*
<eTree>, *717*, *718*
<entry>, *321*, *322*, *324*, *337*, *345*, *2006–2008*
<etym>, *1353*
<event>, *511*, *512*, *521*, *529*, *1357*
<eventName>, *1362*
<explicit>, *1373*
<factuality>, *569*, *1076*, *1381*, *1965*, *2299*
<figure>, *299*, *301*, *302*, *305*, *313*, *314*
<filiation>, *1397*
<floatingText>, *1405*
<forename>, *490*, *492*
<forest>, *715*, *1418*, *1655*
<form>, *322*, *330*, *331*, *335*, *344*, *348–350*, *352*, *354*, *356*, *1063*, *1613*, *1759*, *1873*, *1900*, *2325*
<fs>, *541*, *658*, *664*, *665*, *667*, *668*, *671*, *675*, *678*, *679*, *684*, *687*, *1430*
<fsDecl>, *678–680*, *687*, *1434*, *1436*, *1438*
<fsdLink>, *679*, *1434*, *1440*
<fw>, *457*, *1448*
<geogName>, *498*, *499*, *539*
<gram>, *322*, *332*, *337*, *1486*
<graph>, *691–697*, *699*, *702*, *1492*
<graphic>, *1496*
<head>, *184*, *525*, *601*
<iType>, *331*, *1569*
<ident>, *43*, *1533*
<idno>, *33*, *35*, *144*, *149*, *153*, *155–158*, *307*, *308*, *502*, *531*, *533*, *631*, *632*, *743*, *1356*, *1361*, *1536*, *1661*, *1753*, *1767*, *1825*, *1826*, *1831*, *2109*, *2132*, *2285*
<incipit>, *1548*
<interaction>, *569*, *1076*, *1557*, *1965*, *2299*
<interp>, *1560*
<interpGrp>, *648*, *650*, *651*, *1562*
<join>, *268*
<kinesic>, *306*, *1580*
<lbl>, *344*
<lg>, *168*, *169*, *191*, *232–234*, *237–239*, *241*, *242*, *249*, *265–268*, *904*, *1343*, *1620*, *1644*
<link>, *608*, *617*, *624*
<linkGrp>, *243*, *289*, *291*, *557*, *585*, *604*, *606–609*, *613*, *614*, *652*, *676*, *1629*, *1631*
<list>, *85*, *111*, *120–122*, *125*, *187*, *203*, *221*, *260*, *427*, *739*, *1512*, *1514*, *1586*, *1635*, *1720*, *1993*, *2051*, *2135*, *2185*, *2347*
<listBibl>, *1910*
<listNym>, *1658*
<listPerson>, *85*, *503*, *1356*, *1362*, *1667*, *2148*
<listPlace>, *520*, *521*, *1361*, *1670*
<listRef>, *1910*
<listRelation>, *1677*
<location>, *518*
<m>, *641*, *883*, *1698*
<mapping>, *209*, *214–216*, *1096*, *1098*, *1687*, *1705*, *2361*
<measure>, *111–113*, *386*, *1711*, *2375*
<measureGrp>, *112*, *1714*
<metDecl>, *241*, *245*, *1730*
<mod>, *453*, *460*, *1738*
<move>, *262*, *1762*
<msDesc>, *400*
<name>, *62*, *104–107*, *109*, *148*, *186*, *302*, *307*, *308*, *317*, *318*, *370*, *396*, *485*, *486*, *490*, *495–497*, *499*, *510–512*, *515*, *524*, *570*, *571*, *835*, *836*, *838*, *918*, *947*, *1034*, *1131*, *1207*, *1211*, *1788*, *1848*, *1867*, *1943*, *2130*, *2318*
<node>, *696*, *697*, *1797–1800*
<note>, *118*, *123–125*, *309*, *311*, *344*, *580*, *582*, *583*, *731*, *1640*, *1808*
<num>, *110*, *1817*
<nym>, *538*
<oRef>, *339*, *347*, *1851*
<object>, *534*
<objectName>, *530*, *534*, *1832*
<orgName>, *156*, *158*, *495*, *496*, *524*, *1858*
<origDate>, *531*, *534*, *1826*
<orth>, *349*, *350*, *352*, *1873*
<pause>, *1892*
<pc>, *1897*
<persName>, *309*, *311*, *490*, *494*, *871*, *1910*, *1920*
<phr>, *599*, *638*, *1923*, *1943*
<place>, *517–521*, *524*, *1690*
<placeName>, *517*, *518*, *1690*
<population>, *523*, *1935*, *2210*
<post>, *299*, *302*, *304*, *306*, *314*, *317*, *318*, *838*, *1943–1945*
<preparedness>, *569*, *1076*, *1961*, *1965*, *2299*
<ptr>, *118*, *165*, *302*, *307*, *308*, *314*, *525*, *718*, *893*, *1170*
<purpose>, *569*, *1076*, *1965*, *1988*, *2299*
<rdg>, *464*, *465*, *979*, *1617*
<rdgGrp>, *466–468*, *979*, *2003*
<recording>, *276*, *277*, *2012*, *2014*
<ref>, *63*, *118*, *302*, *314*, *1171*, *1172*
<region>, *498*, *500*, *520*, *525*, *2031*
<relatedItem>, *161–163*, *308*, *2033*
<relation>, *513*, *1362*, *1667*, *1886*, *2036*
<restore>, *435*, *2058*

<rhyme>, *2066*
<roleName>, *493*, *494*
<rs>, *58*, *104–106*, *251*, *490*, *495*, *497*, *500*, *1362*, *1902*, *1903*, *2080*
<s>, *650*
<schemaRef>, *56*, *2100*
<seal>, *394*, *2111*, *2113*
<seg>, *235*, *240*, *286*, *287*, *537*, *598–600*, *622*, *625*, *641*, *974*, *2122*
<settlement>, *498*, *500*, *517*, *935*, *1690*, *1910*, *2142*
<spGrp>, *259*, *269*
, *2180*
<spanGrp>, *646*, *655*, *1640*, *2183*, *2208*
<stage>, *169*, *171*, *250*, *260–262*, *267–269*, *602*, *1070*, *1343*, *1762*, *2052*, *2185*, *2202*
<standOff>, *2208*
<state>, *510*, *514*, *521*, *2214*
<surface>, *415*, *417*, *454*
<surfaceGrp>, *454*
<surname>, *13*, *490*, *491*, *1907*, *2252*
<tag>, *2267*
<tech>, *2278*
<time>, *540–542*
<title>, *38*, *152*, *155*, *158*, *162*, *307*, *308*, *361*, *380*, *381*, *399*, *1019*, *1755*, *1776*, *2033*, *2109*, *2257*, *2316*
<titlePage>, *248*
<titlePart>, *202*, *1296*, *2318*, *2321*
<trait>, *509*, *2332*
<unit>, *113*, *2375*
<unitDef>, *54*, *55*, *859*, *860*, *1711*, *2377*, *2379*
<usg>, *322*, *332*, *335*, *337*, *339*, *341*, *342*, *344*, *345*, *348–350*, *352*, *354*, *851*, *2007*, *2008*, *2127*, *2383*
<valList>, *260*, *505*, *745*, *759*, *781*, *801*, *802*, *2388*, *2392*
<w>, *599*, *641*, *883*, *1698*, *1943*, *1944*
<witDetail>, *2428*
<writing>, *283*
<xr>, *335*, *338*, *343*, *349*, *350*, *352*, *2445*
<typeDesc>, *390*, *2235*, *2342*
<typeNote>, *390*, *2235*, *2342*, *2344*

<u>, *279–284*, *286–291*, *293*, *294*, *306*, *603*, *604*, *621–625*, *655*, *736*, *974*, *2148*, *2183*, *2347*, *2351*
@ulx
<surface>, *407*, *408*, *410*, *411*, *414*, *417–419*, *1378*, *1889*, *2247*, *2249*, *2448*, *2473*
<zone>, *407*, *408*, *410*, *411*, *414*, *417*, *419*, *2448*
@uly
<surface>, *407*, *408*, *410*, *411*, *414*, *417–419*, *1378*, *1889*, *2247*, *2249*, *2448*, *2473*
<zone>, *407*, *408*, *410*, *411*, *414*, *417*, *419*, *2448*
<unclear>, *101*, *102*, *188*, *293*, *294*, *443*, *444*, *592*, *603*, *604*, *2051*, *2349*, *2351*
<undo>, ***451***, *2354*
<unicodeProp>, *210*, *216*, *863*, *1096*, *1098*, *1687*, *2361*
<unihanProp>, *2369*
<unit>, *113*, *1711*, *2374*, *2375*
@unit
<biblScope>, *10*, *35*, *145*, *147–152*, *154*, *159*, *160*, *165*, *878*, *1017*, *1019*, *1232*, *1754*, *1755*, *2132*
<citeStructure>, *140*, *141*, *596*, *597*, *1105*, *1110*
<citedRange>, *1107*
<depth>, *1229*
<dim>, *366*, *534*, *738*, *1243*
<dimensions>, *361*, *366*, *385*, *1243*, *1925*
<gap>, *99*, *100*, *294*, *317*, *318*, *381*, *436*, *443*, *592*, *895*, *1224*, *1453*, *2230*
<height>, *366*, *1243*, *1517*
<measure>, *32*, *111*, *112*, *307*, *386*, *500*, *525*, *886*, *895*, *1375*, *1711*, *1712*
<measureGrp>, *112*, *1714*
<milestone>, *50*, *136–138*, *889*, *1735*
<pc>, *643*
<refState>, *50*, *141*, *2025*
<space>, *455*, *2177*
<timeline>, *288–290*, *605*, *2312*
<unit>, *113*, *2375*
<width>, *366*, *737*, *738*, *1243*, *2423*
<unitDecl>, *54*, *859*, *860*, *1711*, *2374*, *2377*, *2379*
<unitDef>, *54*, *55*, *859*, *860*, *1711*, *2374*, *2377*, *2379*
@unitRef
<measure>, *1711*
<unit>, *2374*
@uri
<equiv>, *746*, *894*, *1348*
@url
<graphic>, *118*, *131*, *213*, *302*, *313*, *314*, *405–408*, *410*, *411*, *413*, *414*, *418*, *419*, *554–556*, *587*, *1378*, *1390*, *1393*, *1483*, *1495*, *1496*, *1805*, *1889*, *2171*, *2247*, *2249*, *2280*, *2448*, *2472*, *2473*
<media>, *1103*, *1717*
<moduleRef>, *773*, *774*, *2103*
<schemaRef>, *56*, *2100*
@usage
<attDef>, *759*, *801*, *802*, *992*, *995*, *1321*
<language>, *58*, *68*, *1601*, *1602*
@use
<citeData>, *141*, *597*, *1105*
<citeStructure>, *140*, *141*, *596*, *597*, *1105*, *1110*
@useSourceRendition
<model>, *761–763*
<usg>, *322*, *332*, *335*, *337*, *339*, *341*, *342*, *344*, *345*, *348–350*, *352*, *354*, *851*, *2007*, *2008*, *2127*, *2383*

<val>, *67*, *68*, *595*, *741*, *758*, *992*, *1314*, *2385*, *2388*
<valDesc>, *758*, *2387*
@valid
<egXML>, *1315*
@validUntil

<dataSpec>, *1233*
<valItem>, *260*, *505*, *745*, *759*, *766*, *767*, *781*, *801*, *802*, *1348*, *1883*, *1884*, *2388*, *2392*
<valList>, *260*, *505*, *745*, *759*, *781*, *801*, *802*, *1348*, *2388*, *2392*
<vAlt>, *670–674*, *682–684*, *687*, *850*, *1386*, *1388*, *1434*, *1438*, *2393*, *2399*, *2407*, *2411*
@value
<age>, *950*
<binary>, *658*, *659*, *662*, *665*, *682*, *683*, *686*, *687*, *1022*, *1024*, *1144*, *1386*, *1388*, *1538*, *1540*, *2305*, *2399*, *2411*
<case>, *1063*
<dataFacet>, *757*, *1193*
<eLeaf>, *1317*
<gen>, *1063*
<gender>, *507*, *508*, *1462*
<gram>, *322*
<localProp>, *209*, *213*, *216*, *1096*, *1483*, *1687*, *2361*
<metSym>, *241*, *244*, *245*, *1730*, *1733*
<mood>, *1759*, *1900*, *2325*
<node>, *699*, *702*
<num>, *110*, *113*, *1817*, *2375*
<number>, *1063*, *1759*, *1900*, *2325*
<numeric>, *314*, *541*, *660*, *669–671*, *1821*
<param>, *764–766*, *1742*, *1745*, *1748*, *1881*
<per>, *1759*, *1900*, *2325*
<persPronouns>, *507*, *1920*
<pos>, *1063*
<sex>, *508*, *1920*, *2145*
<symbol>, *659*, *661*, *663*, *664*, *666–668*, *672–675*, *683*, *684*, *686*, *687*, *847*, *848*, *850*, *1022*, *1144*, *1376*, *1402*, *1430*, *1434*, *1438*, *1445*, *1538*, *1540*, *2208*, *2261*, *2305*, *2393*, *2397*, *2404*, *2405*, *2407*
<tns>, *1759*, *1900*, *2325*
<unicodeProp>, *210*, *216*, *863*, *1096*, *1098*, *1687*, *2361*
<unihanProp>, *2369*
<usg>, *322*
@valueDatcat
<category>, *849*, *851*
<pos>, *353*, *848*, *849*
<symbol>, *661*, *847*
<usg>, *851*
<variantEncoding>, *476*, *478*, *479*, *2395*
@varSeq
<rdg>, *435*, *465*
<rdgGrp>, *468*
<vColl>, *667–673*, *2397*, *2405*
<vDefault>, *682*, *683*, *687*, *1386*, *1388*, *1538*, *2399*, *2411*
@version
<TEI>, *2280*
<application>, *56*, *982*, *984*
<teiCorpus>, *2283*
<transcriptionDesc>, *277*, *2334*
<unicodeProp>, *2361*
<unihanProp>, *2369*
@versionDate
<desc>, *746*, *758*, *759*, *767*, *768*, *894*, *995*, *1198*, *1233*, *1235*, *1741*, *1883*
<gloss>, *758*
<remarks>, *758*
<view>, *270*, *271*, *1058*, *1060*, *2402*
<vLabel>, *666*, *2404*
<vMerge>, *673*, *2405*
<vNot>, *671*, *672*, *674*, *684*, *687*, *2407*
<vocal>, *279*, *281*, *282*, *2410*
<vRange>, *682–684*, *687*, *850*, *1386*, *1388*, *1434*, *1436*, *1438*, *2399*, *2411*

<w>, *313–318*, *413*, *599*, *640–645*, *649–651*, *654*, *655*, *674*, *676*, *723*, *726*, *728*, *882–885*, *974*, *1048*, *1644*, *1698*, *1897*, *1943*, *1944*, *2208*, *2414*
<watermark>, *364*, *385*, *2418*
@weights
<alt>, *453*, *624–626*, *953*, *957*
<when>, *288–291*, *311*, *605*, *2312*, *2420*
@when
<birth>, *488*, *507*, *525*, *570*, *1034*, *1910*
<change>, *67*, *68*, *398*, *459*, *632*, *855*, *971*, *1092*, *1650*, *2010*, *2063*
<conversion>, *55*
<custEvent>, *531*, *1826*
<date>, *38*, *58*, *63–65*, *113*, *114*, *144*, *145*, *147*, *148*, *150*, *155–160*, *162*, *165*, *188*, *251*, *252*, *275*, *277*, *306–308*, *310*, *361*, *399*, *487*, *500*, *511*, *540*, *541*, *544*, *835*, *843–846*, *865*, *966*, *1011*, *1019*, *1043*, *1169*, *1171*, *1172*, *1179*, *1204*, *1205*, *1362*, *1754*, *1920*, *1950*, *1979*, *2012*, *2285*
<death>, *525*, *1211*, *1910*
<docDate>, *46*
<event>, *511*, *521*, *524*, *525*, *529*, *1356*, *1357*, *1652*
<origDate>, *531*, *534*, *1826*
<population>, *523*, *1935*, *2210*
<post>, *311*, *1944*, *1945*
<provenance>, *533*, *1661*, *1825*
<resp>, *154*, *876*
<time>, *114*, *540–542*, *845*, *846*, *2309*
@when-custom
<date>, *543*, *544*, *843*, *844*
@when-iso
<date>, *542*, *844*, *845*
<post>, *299*, *300*, *304*
<time>, *844*
@where
<conversion>, *55*
<event>, *525*, *1356*
<move>, *262*, *1762*
@who
<annotationBlock>, *974*

<change>, *67*, *68*, *632*, *855*, *971*, *1092*, *1650*, *2063*
<kinesic>, *288*, *306*
<move>, *262*, *1762*
<post>, *299–302*, *304–306*, *309–318*, *838*, *1943–1945*
<q>, *619*
<said>, *85*
<setting>, *571*
<sp>, *171*, *257*, *258*, *262*, *833*, *2052*, *2185*
<stage>, *833*
<u>, *279–283*, *286–288*, *290*, *291*, *293*, *294*, *306*, *603*, *604*, *736*, *2148*, *2347*
<vocal>, *279*, *281*, *282*
<writing>, *283*
<width>, *361*, *366*, *385*, *737*, *738*, *1243*, *1714*, *1925*, *2423*
@width
<graphic>, *2472*
<media>, *1717*
<wit>, *470*, *2425*
@wit
<lacunaEnd>, *475*
<lem>, *464–469*, *475*, *477–479*, *483*, *979*, *1590*, *1617*, *2003*, *2428*, *2429*, *2433*
<rdg>, *428*, *463–470*, *472*, *475–481*, *483*, *631*, *979*, *1589*, *1590*, *1617*, *1644*, *2001*, *2003*, *2425*, *2428*, *2429*, *2433*
<witDetail>, *469*, *470*, *2428*
<witDetail>, *469*, *470*, *2428*
<witEnd>, *2429*
@withId
<tagUsage>, *47*, *1792*, *2271*
<witness>, *470–472*, *1683*, *2431*
<witStart>, *475*, *2433*
<writing>, *283*, *2436*
@writtenLines
<layout>, *388*, *1605*

<xenoData>, *66*, *2438–2441*
<xr>, *335*, *338*, *342–344*, *349*, *350*, *352*, *1353*, *1593*, *2445*

<zone>, *407*, *408*, *410*, *411*, *414*, *415*, *417–419*, *447*, *448*, *460*, *1496*, *1626*, *1728*, *2171*, *2448*, *2473*