TEI BY EXAMPLE



MODULE 1: COMMON STRUCTURE, ELEMENTS, AND ATTRIBUTES

Edward Vanhoutte

Ron Van den Branden

Melissa Terras

Centre for Scholarly Editing and Document Studies (CTB) , Royal Academy of Dutch Language and Literature, Belgium, Gent, 9 July 2010

Last updated September 2020

Licensed under a Creative Commons Attribution ShareAlike 3.0 License

TABLE OF CONTENTS

3. Textual Phenomena....1
3.1 Structural Features....2
3.1.1 General....2
3.1.2 Title Pages....3
3.2 Renditional Features....4
3.3 Logical and Semantic Features....5
3.3.1 Referring Strings....6
3.3.2 Dates and Time....7
3.3.3 Numbers and Measures....8
3.3.4 Addresses....9
3.3.5 Abbreviations and Expansions....10
3.4 Analytical Features....11
3.4.1 Notes and Annotations....12
3.4.2 Index Entries....13
3.4.3 Apparent Errors....14
3.4.4 Regularisation and Normalisation....16
3.4.5 Additions, Deletions, and Omissions....17

3. Textual Phenomena

The TEI Guidelines define a set of rules to mark up the phenomena in a wide range of texts and textual objects in a descriptive fashion. Generally speaking, there are four classes of textual phenomena that can be described:

1. Structural
2. Renditional
3. Logical & Semantic
4. Analytic

Structural and renditional features are best understood because they concern a natural kind of textual, though culturally defined, organisation. Books mainly consist of chapters, sections, and paragraphs; poetry is mostly organised in poems, stanzas, and lines; whereas scenes, acts, and parts of speech are structural features of performance texts. In these texts, linguistic units are highlighted by the use of distinct fonts, colours, alignments, italics, underlinings, font weight, etc. These textual codes signal underlying logical and semantic features and functions such as names of organisations, titles of books, distinctive languages, emphatic language use, etc. However, semantic and logical features don't need to be highlighted by means of typographic codes and can occur in texts without any special typographic marking. It needs a thorough understanding of the text and the language to identify them. Semantic and syntactic interpretations added to a text or part of a text that together constitute a new text, we call analytical features. Examples are linguistic (wordclass, morpheme, ...) and narrative (theme, motive, ...) categorisations.

3.1 Structural Features

3.1.1 General

CHALLENGE

Which structural features can commonly be found in prose, verse, and drama?

SOLUTION

- Prose: paragraphs <p>, divisions , headings <head>, lists <list>, listitem <item>, quotations <q>, page breaks <pb>, segments <seg>, figures <figure>, and tables <table>. See Module 3: Prose.
- Verse: line groups <lg> and lines <l>. See Module 4: Poetry.
- Drama: divisions , speeches <sp>, paragraphs <p>, line groups <lg>, lines <l>. and segments <seg>. See Module 5: Drama.

The following example demonstrates a simple use of TEI markup for the encoding of structural features in prose text:

```
<text xmlns="http://www.tei-c.org/ns/1.0">
  <body>
    <div>
      <p>For the first time in twenty-five years, Dr Burt Diddledygook decided not to
       turn up to the annual meeting of the Royal Academy of Whoopledywhaa (RAW). It
       was a sunny day in late September 1960 bang on noontime and Dr Burt was looking
       forward to a stroll in the park instead. He hoped his fellow members of the RAW
       weren't even going to notice his absence.</p>
      <p>Or worse, what would happen when another Academy member had decided to go for a
       stroll in the park instead? He quickly thought up several possible plans:</p>
      <list>
        <item>hide behind a tree and duck</item>
        <item>catch the duck as subject material for a speech on the annual
         meeting</item>
        <item>be frank, meet his colleague, and <list>
```

```
        <item>1. pat him on the shoulder</item>
        <item>2. tell a joke</item>
        <item>3. hand him the duck</item>
        <item>4. offer him a sip from his 2.5 l bottle of coke</item>
        <item>5. pull his beard</item>
      </list></item>
    </list>
    <p>Or maybe he could still announce his absence from the meeting by sending
     an antedated letter of apology to Professor M. Orkelidius, Royal Academy of
     Whoopledywhaa, Queenstreet 81, TB90 00E Whoopledywhaa.</p>
    <p><q>Plenty of options</q>, he thought, sat on a bench and opened the book he
     had taken from the Whoopledywhaaian National Library. It was titled 'While thou
     art here', by Sir Edmund Peckwood. While reading the first sentence, his placid
     expression turned to a certain je ne sais quoi: <q>For the first time in twenty
    -five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of
     the Royal Academy of Whoopledywhaa.</q></p>
   </div>
  </body>
</text>
```

Example 10. Encoding structural features in prose text.

3.1.2 Title Pages

Title pages may be encoded within <front> or <back> by using the element <titlePage>.[1] A title page commonly contains the title of the work (<docTitle>), which can consist of several subsections or divisions (<titlePart>), with an @type attribute documenting their role. The name of the author of the document (<docAuthor>) often occurs inside a byline (<byline>), which contains the primary statement of responsibility given for a work. Other components of <titlePage> may be the edition statement (<docEdition>), the date of a document (<docDate>), and the imprint statement (<docImprint>), which may further contain the place of publication (<pubPlace>), a date (<date> or <docDate>), and names (<name>) of, e.g., the publisher (<publisher>). Besides

.....

1 <titlePage> must not be confused with <fileDesc>, which may contain <titleStmt> and <publicationStmt>. Whereas <titlePage> is used for the transcription and encoding of the physical title page in <text>, <fileDesc> is part of the <teiHeader> section containing meta-information, in this case a bibliographic description of the electronic file.

this information, a <titlePage> may also contain an anonymous or attributed quotation (<epigraph>), a formal statement authorizing the publication of a work (<imprimatur>), and/or an inline graphic, illustration, or figure (<graphic>).

```
<front xmlns="http://www.tei-c.org/ns/1.0">
  <titlePage>
    <docAuthor>Roy Offire</docAuthor>
    <docTitle>
      <titlePart type="main">The Strange Adventures of Dr. Burt Diddledygook</titlePart>
      <titlePart type="sub">Wanderings in the life of a buoyant academic</titlePart>
    </docTitle>
    <byline>Transcribed from the diaries.</byline>
    <docEdition>First Edition</docEdition>
    <docImprint><pubPlace>Kirkcaldy</pubPlace>, <publisher>Bucket
     Books</publisher>, <docDate>1972</docDate></docImprint>
  </titlePage>
</front>
```

Example 11. Encoding a title page with <titlePage>.

3.2 Renditional Features

Some textual features are commonly rendered in a text using some kind of highlighting. The TEI Guidelines define highlighting as "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" (TEI Guidelines, section 3.3.1 What Is Highlighting?). If the encoder prefers only to signal this highlighting, and not the underlying reason, the generic element <hi> (highlighting) can be used with a @rend or @rendition attribute describing its appearance in the text. Since these attributes may need to express a wide range of typographic features, no formal values are being defined by the TEI Guidelines: encoders should device their own value system.

```
<p xmlns="http://www.tei-c.org/ns/1.0">For the first time in twenty-five
 years, Dr Burt Diddledygook decided not to turn up to the annual meeting of
 the <hi rend="italic">Royal Academy of Whoopledywhaa</hi> (RAW). It was a sunny day
```

```
in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in
the park instead. He hoped his fellow members of the RAW weren't even going to notice
his absence.</p>
```

Example 12. Encoding typograhpically marked text without indicating the underying meaning, with `<hi>`.

Encoders, however, commonly prefer to indicate the reason underlying the highlighting by documenting logical or semantic information about the highlighted word or phrase. Where possible, this can be done using the elements discussed in the following sections.

3.3 Logical and Semantic Features

Highlighted words or phrases in a text are commonly distinguished from their surroundings for a reason. Only a thorough understanding of the text and the language can lead to a correct identification and interpretation. The underlying semantics may be encoded with more specific elements than the generic <hi> element. Highlighting is commonly used to render the following logical and semantic features:

- Emphasis (<emph>), foreign words (<foreign>), and other linguistically distinct uses (<distinct>) of highlighting.
- The use of quotation marks in the representation of speech and thought (<said>), quotation (<quote>), cited quotation (<cit>), words or phrases mentioned (<mentioned>), and words or phrases for which the author or narrator indicates a disclaiming of responsibility (<soCalled>). See Module 3: Prose, section 2.5.
- Technical terms (<term>), glosses (<gloss>), or documentation of XML elements, attributes, and classes with <altIdent>, <desc>, <equiv>. See Module 3: Prose, section 2.5, and Module 8: Customising TEI, ODD, Roma.

```
<p xmlns="http://www.tei-c.org/ns/1.0"><q>Plenty of options</q>, he thought, sat
 on a bench and opened the book he had taken from the Whoopledywhaaian National
 Library. It was titled 'While thou art here', by Sir Edmund Peckwood. While reading
 the first sentence, his placid expression turned to a certain <foreign>je ne
 sais quoi</foreign>: <quote>For the first time in twenty-five years, Dr Burt
 Diddledygook decided not to turn up to the annual meeting of the Royal Academy of
 Whoopledywhaa.</quote></p>
```

Example 13. Encoding the reason for text highlighting with specific TEI elements.

However, words or phrases carrying semantic and logical information don't need to be highlighted by means of typographic codes and can occur in texts unmarked. Think about titles (<title>), names (<name>), numbers (<num>), measures (<measure>), dates (<date>), addresses (<address>), and abbreviations (<abbr>).

3.3.1 Referring Strings

Proper nouns name people, places, and objects and are easily traceable in a text, since they commonly appear with the first letter in upper case. This may be encoded with <name> carrying a @type attribute specifying the kind of object referred to.

```
<p xmlns="http://www.tei-c.org/ns/1.0">'Plenty of options', he thought, sat on
 a bench and opened the book he had taken from the Whoopledywhaaian National
 Library. It was titled 'While thou art here', by Sir <name type="person">Edmund
 Peckwood.</name> While reading the first sentence, his placid expression
 turned to a certain je ne sais quoi: 'For the first time in twenty-five years,
 Dr <name type="person">Burt Diddledygook</name> decided not to turn up to the annual
 meeting of the <name type="organisation">Royal Academy of Whoopledywhaa</name>.'</p>
```

Example 14. Encoding proper names with the <name> element.

However, people, places, and objects may also be referred to with common nouns, for which the element <rs> (referring string) may be used. This element may also carry a @type attribute specifying the kind of object referred to.[2]

```
<p xmlns="http://www.tei-c.org/ns/1.0">'Plenty of options', <rs type="person">he</rs>
 thought, sat on a bench and opened the book <rs type="person">he</rs> had taken
 from the <rs type="organisation">Whoopledywhaaian National Library</rs>. It was
 titled 'While thou art here', by Sir <name type="person">Edmund Peckwood.</name>
 While reading the first sentence, <rs type="person">his</rs> placid expression
```

.....

2 The <rs> element may be used for any reference to a person, place or object in the form of a proper noun, a noun phrase, or a common noun. The <name> element may be used synonymously with the <rs> element in the special cases of referencing strings which consist only of proper nouns. The choice between <rs> or <name> in these cases is the encoder's. Both elements can nest: for example, <name> may also nest inside <rs> where a proper name is part of a larger referring string, as in `<rs type="organisation">Royal Academy of <name type="place">Whoopledywhaa</name></rs>`

```
turned to a certain je ne sais quoi: 'For the first time in twenty-five years,
Dr <name type="person">Burt Diddledygook</name> decided not to turn up to the annual
meeting of the <name type="organisation">Royal Academy of Whoopledywhaa</name>.'</p>
```

Example 15. Encoding referring strings with `<rs>`.

3.3.2 Dates and Time

Any expression defining a date or time may be encoded with the corresponding elements <date> and <time>. The system or calendar to which the date belongs may be documented using a @calendar attribute. The @when attribute supplies the value of a date or time in a standard form, which is useful for text processing.

The normalised representation of the content of the <date> element should conform to a valid W3C schema datatype for expressing temporal data:

- `<date when="2009" calendar="Gregorian">2009</date>`
- `<date when="2009-12" calendar="Gregorian">December 2009</date>`
- `<date when="2009-12-31" calendar="Gregorian">31 Dec 2009</date>`
- `<date when="2009-12-31" calendar="Gregorian">New Year's Eve 2009</date>`
- `<date when="2009-12-31" calendar="Persian">Panjshanbeh 10 Dey 1388</date>`
- `<date when="--12-31">last day of December</date>`
- `<date when="--12">December</date>`
- `<date when="---31">thirty-first of the month</date>`

The same counts for the normalized representation of the content of <time>:

- `<time when="23:55:00">11:55 pm</time>`
- `<time when="23:55:00">five minutes before midnight</time>`
- `<time when="2009-12-31T23:55:00">five minutes before the new year 2010</time>`[3]

The <date> element can also be used to mark a span of time, using the @from and @to attributes, or a range of time, using the @notBefore and @notAfter attributes:

.....

3 The last example also includes a date string and can equally well be tagged as `<date when="2009-12-31T23:55:00">five minutes before the new year 2010</date>`.

```
<p xmlns="http://www.tei-c.org/ns/1.0">For the first time
 in <date from="1935" to="1960">twenty-five years</date>, Dr Burt Diddledygook decided
 not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa (RAW). It
 was a sunny day in <date notBefore="1960-09-15" notAfter="1960-09-30">late September
 1960</date> bang on <time when="12:00:00">noontime</time> and Dr Burt was looking
 forward to a stroll in the park instead. He hoped his fellow members of the RAW weren
't even going to notice his absence.</p>
```

Example 16. Expressing more dating nuances with `@from`, `@to`, `@notBefore`, and `@notAfter`.

In this example, the @from and @to attributes in the first <date> element express a period of time, spanning from 1935 to 1960. In the second <date> element, the combination of @notBefore and @notAfter indicates a time range in the second half of September 1960.

> REFERENCE
>
> See section 13.1.2 Dating Attributes of the TEI Guidelines for a comprehensive explanation of the use and combinations of these dating attributes.

3.3.3 Numbers and Measures

Numbers and measures may be encoded using <num> and <measure> respectively.

<num> may contain numbers, written in any form. The attribute @type can be used to indicate the type of numeric value, and @value to supply the value of the number in standard form.

```
<p xmlns="http://www.tei-c.org/ns/1.0">For the first time
 in <num type="cardinal" value="25">twenty-five</num> years, Dr Burt Diddledygook
 decided not to turn up to the annual meeting of the Royal Academy of
 Whoopledywhaa.</p>
```

Example 17. Encoding numbers with `<num>`.

Here are more examples of the standardisation of numbers:

- `<num value="25">xxv</num>`

- `<num type="percentage" value="25">twenty-five percent</num>`
- `<num type="percentage" value="25">25%</num>`
- `<num type="ordinal" value="25">25th</num>`

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 to be present in every case. These three components may be encoded on a <measure> element with the attributes @quantity, @unit, and @commodity.

```
<p xmlns="http://www.tei-c.org/ns/1.0">Or worse, what would happen when another Academy
 member had decided to go for a stroll in the park instead? He quickly thought up
 several possible plans:</p>
<list xmlns="http://www.tei-c.org/ns/1.0">
  <item>hide behind a tree and duck</item>
  <item>catch the duck as subject material for a speech on the annual meeting</item>
  <item>be frank, meet his colleague, and <list>
    <item>1. pat him on the shoulder</item>
    <item>2. tell a joke</item>
    <item>3. hand him the duck</item>
    <item>4. offer him a sip from his <measure type="volume" quantity="2
    .5" unit="litre" commodity="coca-cola">2.5 l bottle of coke</measure></item>
    <item>5. pull his beard</item>
  </list></item>
</list>
```

Example 18. Encoding measures with `<measure>`.

3.3.4 Addresses

E-mail addresses can be encoded with the <email> element.

```
<email xmlns="http://www.tei-c.org/ns/1.0">M.Orkelidius@raw.org</email>
```

Example 19. Encoding an e-mail address with `<email>`.

A postal address can be encoded with the <address> element. It can contain a number of <addrLine> elements, one for each address line.

```
<p xmlns="http://www.tei-c.org/ns/1.0">Or maybe he could still announce his absence
 from the meeting by sending an antedated letter of apology to <address>
  <addrLine>Professor M. Orkelidius</addrLine>
  <addrLine>Royal Academy of Whoopledywhaa</addrLine>
  <addrLine>Queenstreet 81</addrLine>
  <addrLine>TB90 00E Whoopledywhaa</addrLine>
</address></p>
```

Example 20. Encoding a postal address with <address>.

Alternatively, an address can be encoded in more detail, with more semantically rich elements such as <street>, <postCode> and <postBox>. Names of people, organisations, companies, etc. may be encoded using <name>, with a @type attribute indicating the type of object which is being named.

```
<p xmlns="http://www.tei-c.org/ns/1.0">Or maybe he could still announce his absence
 from the meeting by sending an antedated letter of apology to <address>
  <name type="person">Professor M. Orkelidius</name>
  <name type="organisation">Royal Academy of Whoopledywhaa</name>
  <street>Queenstreet 81</street>
  <postCode>TB90 00E</postCode>
  <name type="city">Whoopledywhaa</name>
</address></p>
```

Example 21. Encoding the components of a postall address with specific elements.

3.3.5 Abbreviations and Expansions

It is sometimes useful to encode abbreviations and their expansions in texts. This facilitates special processing, regularisation by the full form of an abbreviation, or the rendering of different possible expansions of an abbreviation. Abbreviations may be marked using <abbr>. The @type attribute may be used to distinguish types of abbreviations by their function:

```
<p xmlns="http://www.tei-c.org/ns/1.0">For the first time in twenty-five
 years, <abbr type="title">Dr</abbr> Burt Diddledygook decided not to turn up to
 the annual meeting of the Royal Academy of Whoopledywhaa (<abbr type="acronym">R
AW</abbr>). It was a sunny day in late September 1960 bang on noontime
 and <abbr>Dr</abbr> Burt was looking forward to a stroll in the park instead. He hoped
 his fellow members of the <abbr type="acronym">RAW</abbr> weren't even going to notice
 his absence.</p>
```

Example 22. Encoding abbreviations with `<abbr>`.

Alternatively, and depending on the encoder's preference, the expansion of an abbreviation may be encoded with <expan>. This is often done when the editor or encoder of a text has silently expanded the abbreviation for whatever reason. It is equally possible to record both the (original) abbreviation and the (editorial) expansion by wrapping both in a <choice> element.

```
<p xmlns="http://www.tei-c.org/ns/1.0">For the first time in twenty-five
 years, <choice>
  <abbr type="title">Dr</abbr>
  <expan>Doctor</expan>
</choice> Burt Diddledygook decided not to turn up to the annual meeting of the Royal
 Academy of Whoopledywhaa (<choice>
  <abbr type="acronym">RAW</abbr>
  <expan>Royal Academy of Whoopledywhaa</expan>
</choice>).</p>
```

Example 23. Combining both abbreviations and their expansions in `<choice>`.

3.4 Analytical Features

The analysis of texts can generate information which may be added to the text and encoded as metadata or as part of the text. Explicit notes are the most common example of the latter while editorial statements like correction of errors, regularisation of spelling variants, or the marking of the text for indexing purposes are examples of the former. The creation of index entries also enhances further analysis of the text.

3.4.1 Notes and Annotations

The most explicit form of textual annotation is the addition of notes to the text using <note>. This element serves for the encoding of all kinds of annotations, whether they are already present in the text or supplied by the editor; 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. The @type attribute can be used to distinguish between different types of annotations. In a @resp attribute, the person or other agency responsible for the content of the note can be identified, pointing to the @xml:id value of an element that identifies this person or agency. Where possible, a note can be inserted in the text at the point where its identifier or mark first appears. The location of the note may be documented using a @place attribute.

```
<p xmlns="http://www.tei-c.org/ns/1.0">'Plenty of options', he thought, sat on
 a bench and opened the book he had taken from the Whoopledywhaaian National
 Library<note n="1" place="foot" type="authorial">The National Library of
 Whoopledywhaa was founded in 1886 with the acquisition of the library of the
 late King Anthony.</note>. It was titled 'While thou art here', by Sir Edmund
 Peckwood<note type="editorial" resp="#EV">The manuscript reads 'Petwood'.</note>.
 While reading the first sentence, his placid expression turned to a certain je ne sais
 quoi: 'For the first time in twenty-five years, Dr Burt Diddledygook decided not to
 turn up to the annual meeting of the Royal Academy of Whoopledywhaa.'</p>
```

Example 24. Encoding an editorial annotation with <note>.

Here, an editorial annotation is inserted into the text, using the <note> element. Its @type attribute indicates it is an "editorial" annotation; the person responsible for its content is pointed to with the @resp attribute. In this case, it is referring to another element in the same document, with @xml:id="EV". In the @resp attribute, this ID value is preceded with a hash character (#), in order to indicate it as the identifier part of a formal URI reference.

REFERENCE

See section 3.6. Simple Links and Cross-References in the TEI Guidelines for a full discussion of notes which are encoded not at the point of attachment but at the point of appearance, e.g., at the end of a chapter or a volume. See chapter 16. Linking, Segmentation, and Alignment for mechanisms to encode multiple views of larger or heterogeneous spans of text. See section 17.3. Spans and Interpretation for a discussion of advanced interpretive annotations.

3.4.2 Index Entries

Pre-existing indexes may be encoded as plain lists (<list>) inside in the <front> or <back> sections of a <text>, for example. On the other hand, in order to generate new indexes from machine readable text, the location to be indexed can be marked with an <index> element. When the text is being indexed on multiple levels, the name of the index can be given in an @indexName attribute. The term to be indexed should appear in a <term> element inside <index>.

```
<p xmlns="http://www.tei-c.org/ns/1.0">'Plenty of options', he thought, sat on
 a bench and opened the book he had taken from the Whoopledywhaaian National
 Library<index indexName="institutions">
  <term>Library</term>
  <index>
    <term>National</term>
  </index>
</index>. It was titled 'While thou art here', by Sir Edmund Peckwood. While
 reading the first sentence, his placid expression turned to a certain je ne
 sais quoi: 'For the first time in twenty-five years, Dr Burt Diddledygook
 decided not to turn up to the annual meeting of the Royal Academy of
 Whoopledywhaaw<index indexName="institutions">
  <term>Academy</term>
```

```
  <index>
    <term>Royal</term>
  </index>
</index>.'</p>
```

Example 25. Encoding index entries with `<index>`.

Notice, how <index> entries can nest in order to create multi-level index entries. With this encoding in place, it will be possible to create an "institutions" index, with the terms "Library, National" and "Academy, Royal," referencing the location of the original <index> element in the text.

REFERENCE

See section 3.8.3 Index Entries of the TEI Guidelines for a full discussion of the TEI encoding strategies applied to indexes.

3.4.3 Apparent Errors

Apparent errors in the text may be indicated using the <sic> element, or corrected inside <corr>.

```
<p xmlns="http://www.tei-c.org/ns/1.0">It was titled 'While thou art here', by Sir
 Edmund <sic>Petwood</sic>.</p>
```

```
<p xmlns="http://www.tei-c.org/ns/1.0">It was titled 'While thou art here', by Sir
 Edmund <corr>Peckwood</corr>.</p>
```

Example 26. Alternative encodings: an apparent error with `<sic>`, or its correction with `<corr>`.

Alternatively, the encoder may both record the original source text and provide a correction by using both <sic> and <corr> (in either order) wrapped in a <choice> element.

```
<p xmlns="http://www.tei-c.org/ns/1.0">It was titled 'While thou art here', by Sir
 Edmund <choice>
  <corr>Peckwood</corr>
  <sic>Petwood</sic>
```

```
</choice>.</p>
```

Example 27. Combining both errors and their corrections in `<choice>`.

The encoder may encode the degree of certainty associated with the intervention or interpretation using a @cert attribute, and indicate the agency responsible for the intervention or interpretation (for instance an editor or transcriber), using @resp. The value of @resp is a pointer to an element in the document header that is associated with a person responsible for the intervention.

```
<p xmlns="http://www.tei-c.org/ns/1.0">It was titled 'While thou art here', by Sir
 Edmund <choice>
  <corr cert="high" resp="#EV">Peckwood</corr>
  <sic>Petwood</sic>
</choice>.</p>
```

Example 28. Identifying the person responsible for a correction with `@resp`, and indicating a degree of certainty with `@cert`.

The attribute value `"#EV"` points to a <name> element in the <teiHeader>, for example in the <respStmt> section:

```
<respStmt xmlns="http://www.tei-c.org/ns/1.0">
  <resp>editor</resp>
  <name xml:id="EV">Edward Vanhoutte</name>
</respStmt>
```

Example 29. Identifying an editor for the electronic text with `<respStmt>`.

REFERENCE

See Module 6: Primary Sources, section 4 for a fuller treatment of editorial interventions.

3.4.4 Regularisation and Normalisation

Standard or regularised forms for variant forms or non-standard spelling may be provided for a number of reasons. This is called regularisation or normalisation. The original, non-normalized form may be flagged using the <orig> element.

```
<p xmlns="http://www.tei-c.org/ns/1.0">It was titled
 'While <orig>thou</orig> <orig>art</orig> here', by Sir Edmund Peckwood.</p>
```

Example 30. Explicitly encoding a word as an original form in the source text, with `<orig>`.

If the encoder wants to indicate that certain words have been normalised, which means modernisation of spelling in this example, the <reg> element may be used.

```
<p xmlns="http://www.tei-c.org/ns/1.0">It was titled
 'While <reg>you</reg> <reg>are</reg> here', by Sir Edmund Peckwood.</p>
```

Example 31. Encoding a regularised form with `<reg>`.

Alternatively, the encoder may decide to record both the original form <orig> and the regularised form <reg>, wrapped inside a <choice>. In the case of the modernisation of spelling, an electronic text could thus serve as the basis of an old- or new-spelling edition.

```
<p xmlns="http://www.tei-c.org/ns/1.0">It was titled 'While <choice>
  <orig>thou</orig>
  <reg>you</reg>
</choice> <choice>
  <orig>art</orig>
  <reg>are</reg>
</choice> here', by Sir Edmund Peckwood.</p>
```

Example 32. Combining both original forms and their regularisations in `<choice>`.

The @resp attribute may be used to specify the person or agency responsible for the regularisation or normalisation.

```
<p xmlns="http://www.tei-c.org/ns/1.0">It was titled 'While <choice>
  <orig>thou</orig>
  <reg resp="#EV">you</reg>
</choice> <choice>
  <orig>art</orig>
  <reg resp="#EV">are</reg>
</choice> here', by Sir Edmund Peckwood.</p>
```

Example 33. Identifying the person responsible for a normalisation with `@resp`.

3.4.5 Additions, Deletions, and Omissions

Another editorial intervention in the text may be the documentation and creation of additions, deletions, and omissions. When transcribing a source document, <gap> may be used to indicate a point where material has been omitted, both because the material is illegible, invisible, or inaudible in the source, and because the editor or transcriber has decided to omit material for editorial reasons or as part of sampling practice. The reason for omission may be given in a @reason attribute. Sample values include `"sampling"`, `"illegible"`, `"inaudible"`, `"irrelevant"`, `"cancelled"`. Additional attributes like @extent and @unit may document the amount of characters, words, lines or any other unit omitted.[8]

```
<p xmlns="http://www.tei-c.org/ns/1.0">For the first time in twenty-five years, Dr
 Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy
 of Whoopledywhaa (RAW). <gap reason="irrelevant" unit="words" extent="32"/>It was a
 sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a
 stroll in the park instead. He hoped his fellow members of the RAW weren't even going
 to notice his absence.</p>
```

Example 34. Encoding omitted text with `<gap>`.

.....

8 If the omission is an editorial policy decision, e.g., the systematic exclusion of marginal commentaries from an encoding, the full details of the policy should be documented in <editorialDecl> inside the <encodingDesc> of the TEI Header. See Module 2: The TEI Header, section 3.2.

The <gap> element may appear as an empty element, but my also contain a description of the material omitted using <desc>.

```
<p xmlns="http://www.tei-c.org/ns/1.0">For the first time in twenty-five years, Dr
 Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of
 Whoopledywhaa (RAW). <gap reason="irrelevant" unit="words" extent="32">
  <desc>Commentary on the founding charter of the RAW</desc>
</gap> It was a sunny day in late September 1960 bang on noontime and Dr Burt was
 looking forward to a stroll in the park instead. He hoped his fellow members of the R
AW weren't even going to notice his absence.</p>
```

Example 35. Describing text fragments omitted with <desc> inside <gap>.

Where words or phrases of moderate lengths have been added or deleted in the source text, this may be recorded using <add> and <del>. As with all TEI elements, information on the actual rendition of the additions and deletions can be provided in the global @rend attribute. Additionally, the place of the addition may also be recorded using @place. See Module 6: Primary Sources, section 3.1.1 for a detailed discussion of these elements and their attributes.[9]

```
<p xmlns="http://www.tei-c.org/ns/1.0">For the first time in twenty-five years, Dr
 Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy
 of Whoopledywhaa (RAW). It was a sunny day in <add place="supralinear">late</add>
 September 1960 bang on noontime and Dr Burt was looking forward to
 a <del rend="overstrike">walk</del><add place="infralinear">stroll</add> in the park
 instead. He hoped his fellow members of the RAW weren't even going to notice his
 absence.</p>
```

Example 36. Encoding "authorial" additions and deletions in the source text with <add> and <del>, respectively.

When additions and deletions can be considered a single intervention in the text, <add> and <del> can be grouped inside <subst> (substition).

.....

9 When an editor wants to mark his or her own additions as editorial interventions in the text, <corr> or <supplied> should be used, not <add>. See Module 6: Primary Sources, section 4. For longer additions and deletions, <addSpan> and <delSpan> may be used. See Module 6: Primary Sources, section 3.1.2.

```
<p xmlns="http://www.tei-c.org/ns/1.0">For the first time in twenty-five years, Dr
 Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy
 of Whoopledywhaa (RAW). It was a sunny day in <add place="supralinear">late</add>
 September 1960 bang on noontime and Dr Burt was looking forward to a <subst>
  <del rend="overstrike">walk</del>
  <add place="infralinear">stroll</add>
</subst> in the park instead. He hoped his fellow members of the RAW weren't even going
 to notice his absence.</p>
```

Example 37. Grouping related additions and deletions that together make up one substitution in the text, inside `<subst>`.

Where deletions in the source text cannot be read with confidence, <unclear> should be used with the @reason attribute indicating that the difficulty of transcription is due to deletion. See Module 6: Primary Sources, section 4.1.