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

5. Global Attributes....1

5.1 @xml:id....2

5.2 @n....4

5.3 @xml:lang....5

5.4 @xml:base....6

5.5 @xml:space....6

5.6 @rend....7

5.7 @style....8

5.8 @rendition....9

5.9 @cert....10

5.10 @resp....11

5.11 @source....11

5. Global Attributes

Just as any XML element, TEI elements can carry one or more attributes which provide additional information, and function as their qualifiers and quantifiers. The full list of all attributes defined in TEI is available as Appendix D Attributes of the TEI Guidelines. A couple of these attributes can occur on all TEI elements: those are defined as "global attributes," in the att.global attribute class, and its subclasses. Not all of those subclasses are always present for all TEI documents (see Module 8: Customising TEI, ODD, Roma, section 5.1 for more information on including TEI modules in a TEI schema), but a number of attribute classes are *always* present in any TEI schema (since they are defined in the **`tei`** module). Together, they define 11 global attributes, available on any TEI element:

att.global

`@xml:id`

provides a unique identifier for an element.

`@n`

provides a number or other label for an element, which does not need to be unique within the document.

`@xml:lang`

indicates the language of an element using a "tag" generated according to BCP 47.

`@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.

att.global.rendition

@rend

indicates how the element was rendered or presented in the source text.

@style

contains an expression in some formal style definition language which defines the rendering or presentation used for this element in the source text.

@rendition

points to a description of the rendering or presentation used for this element in the source text.

att.global.responsibility

@cert

signifies the degree of certainty associated with the intervention or interpretation

@resp

indicates the agency responsible for the intervention or interpretation

att.global.source

@source

specifies the source from which some aspect of this element is drawn.

5.1 `@xml:id`

The `@xml:id` attribute provides a unique identifier for the element bearing the attribute. The identifier must be unique in the whole XML document. If there is another element in the XML document bearing the same identifier as a value for this attribute, a validating XML parser will signal a syntax error. Conforming to the World Wide Web Consortium's XML Recommendations, the attribute value must be a legal name, which means

that it must start 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. The use of the colon in a unique identifier is forbidden as it has the specific purpose of indicating namespace prefixes in XML.

CHALLENGE

Which one of the following examples demonstrates a correct use of @xml:id and why?

1.

```
<p xml:id="p:1">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>
<p xml:id="p:2">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.</p>
<p xml:id="p:2">He hoped his fellow members of the Royal Academy weren't even going to notice his absence.</p>
```

2.

```
<p xml:id="1">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>
<p xml:id="2">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.</p>
<p xml:id="3">He hoped his fellow members of the Royal Academy weren't even going to notice his absence.</p>
```

3.

```
<p xml:id="p1">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>
<p xml:id="p2">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.</p>
<p xml:id="p3">He hoped his fellow members of the Royal Academy weren't even
 going to notice his absence.</p>
```

SOLUTION

Example 3 demonstrates a correct use of the @xml:id attribute: the attribute value is unique, it starts with a letter, and contains no illegal characters. The attribute values in example 1 are not unique and the use of a colon is forbidden. The attribute values in example 2 start with a number, which is not allowed.

5.2 @n

The @n attribute also provides an identifier for an element, but its value doesn't need to be a legal XML name. This means that they don't have to be unique inside the XML document and they may start with and contain any character. Typically @n is used to number or label elements. All @n values in the following examples are legal:

```
<p n="1">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>
<p n="p2">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.</p>
<p n="paraghraph 3">He hoped his fellow members of the Royal Academy weren't even going
 to notice his absence.</p>
```

Example 42. Providing informal labels for elements in the @n attribute.

Although by no means mandatory, it often makes sense to enrich the structural units of a document (e.g., lines in a poem) with some sort of identification (in @xml:id) or reference mechanism (in @n). Of course, when dealing with complex and/or long documents, this labelling could become a rather demanding task in itself. Fortunately, this job can be done automatically by an XML processor, which can identify the sequential position of one element

within another in an XML document without any additional tagging. Instead of manually providing mechanical references for a long poem or collection of poems, you could as well instruct an XML processor to either enrich the TEI encoding and add @xml:id or @n attributes with appropriate values, or to automatically deduct such reference systems from your markup and present them while rendering the document (e.g., in an HTML version of a poem).

REFERENCE

See section 3.10.2 Creating New Reference Systems of the TEI Guidelines for guidance on creating sensible reference systems for text structures.

5.3 @xml:lang

The language of the content of a given element may be documented as the value of an @xml:lang attribute. If it is not specified, the value is inherited from that of the immediately enclosing element. Therefore, it is simplest to specify the base language of a text on the <TEI> element and override that with @xml:lang attributes only for those elements with a different language.

```
<p xmlns="http://www.tei-c.org/ns/1.0" xml:lang="en">'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 Edmund
 Peckwood. While reading the first sentence, his placid expression turned to a
 certain <foreign xml:lang="fr">je ne sais quoi</foreign>: '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 43. Specifying the language of the content of an element with the @xml:lang attribute.

REFERENCE

The values for the @xml:lang attribute must be constructed in a uniform way as explained in section vi.1. Language identification of the TEI Guidelines.

5.4 `@xml:base`

Many TEI attributes take a URI reference as their value. Those can be either *absolute* (starting with the protocol, such as `http:`, `ftp:`, ...) or *relative* (either starting with a local file name, such as `names.xml`, and/or a fragment identifier, such as `#EV`). The `@xml:base` attribute can be used to set a context for all relative URLs appearing within the element on which the `@xml:base` attribute is specified. For example:

```
<p xmlns="http://www.tei-c.org/ns/1.0" xml:base="../xml/">'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="names.xml#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 44. Providing a context for the resolution of relative URIs with the `@xml:base` attribute.

In this example, the relative URI `names.xml#EV` will be resolved to a subfolder named **`xml`** of the folder containing the electronic text containing that reference. Hence, the URI reference will be evaluated as `../xml/names.xml#EV`.

5.5 `@xml:space`

This global attribute provides a mechanism for indicating to systems processing an XML file how they should treat white space. It has two possible values: `"default"` (white space will most probably be normalised during processing) and `"preserve"` (white space should be preserved as is during processing).

```
<choice xmlns="http://www.tei-c.org/ns/1.0">
  <sic xml:space="preserve">white     space</sic>
  <corr>white space</corr>
</choice>
```

Example 45. Specifying how white space should be handled during processing with `@xml:space`.

In this example, the `@xml:space` on the <sic> element specifies that the (unusual) spacing in the original form should be preserved when this document is being processed.

Notice, how the `@xml:space` attribute is rarely used in TEI documents because such layout features are generally expressed more confidently, and descriptively, with TEI elements such as <lb> or <space>, or using the renditional attributes described next.

5.6 `@rend`

The `@rend` attribute is used to document information about the physical appearance of the text in the source. In the following example, it is used to indicate that the title, the French phrase, and the name of the Royal Academy are printed in italics:

```
<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 <title type="m" rend="italics">While thou art here</title>, by
 Sir Edmund Peckwood. While reading the first sentence, his placid expression
 turned to a certain <hi rend="italics">je ne sais quoi</hi>: 'For the first
 time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the
 annual meeting of the <name type="organisation" rend="italics">Royal Academy of
 Whoopledywhaa</name>.'</p>
```

Example 46. Indicating specific renditional features in the source text with `@rend`.

The value for `@rend` can take the form of a white space separated list of idiosyncratic keywords, which an XML processor can act upon when rendering the document. This means that multiple renditional features can be enumerated with `@rend`.

5.7 @style

The @style attribute can also be used to document information about the physical appearance of the text in the source. Contrary to @rend, @style must express this information in some formal style definition language. This will most often be CSS, although others are possible as well. The name of that formal style definition language can be given in the <encodingDesc> section of the header, in a <styleDefDecl> element:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <!-- ... -->
    <encodingDesc>
      <styleDefDecl scheme="css" schemeVersion="2.1"/>
      <!-- ... -->
    </encodingDesc>
    <!-- ... -->
  </teiHeader>
  <text>
    <body>
      <!-- ... -->
      <p>'Plenty of options', he thought, sat on a bench and opened the
       book he had taken from the Whoopledywhaaian National Library. It was
       titled <title type="m" style="font-style:italic;">While thou art here</title>,
       by Sir Edmund Peckwood. While reading the first sentence, his placid expression
       turned to a certain <hi rend="italics">je ne sais quoi</hi>: 'For the first
       time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the
       annual meeting of the <name type="organisation" style="font-weight:italic;">Royal
       Academy of Whoopledywhaa</name>.'</p>
      <!-- ... -->
    </body>
  </text>
</TEI>
```

Example 47. Indicating specific renditional features in the source text with @rend.

5.8 @rendition

Whereas the @rend and @style attributes documents the appearance of text locally, i.e., attached to an element, the @rendition attribute points to a description of the rendering or appearance in the header (<teiHeader>), more specifically inside a <tagsDecl> inside the <encodingDesc> section. This is done in free text or using a formal language inside a <rendition> element. This way, only one description of the rendering must be given, which can be referred to with @rendition attributes on elements in the text. The advantage of this system becomes clear when both @rendition and @rend are used for occurrences of a given element. While the former refers to an overall description of the appearance of that element in the source, the latter documents the local deviation from that generally imposed rendition.

In the following example, we see a description of the overall rendering of <hi> elements in a document, in the <tagsDecl> element inside the <encodingDesc> section of <teiHeader>. The @gi attribute of <tagUsage> names the elements for which the rendition described in <rendition> is documented. The formal namespace in which the tags described in <tagUsage> are defined, must be specified in the @name attribute of a surrounding <namespace> element. The value of the @rendition attribute of <tagUsage> refers to <rendition> by way of the latter's @xml:id attribute. This way, all <hi> elements inside <text> have the style, defined as "italic", as their default rendition. In the following example, the third occurrence of the <hi> element in the text documents a deviant rendition, by means of the @rend attribute.

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <!-- ... -->
    <encodingDesc>
      <!-- ... -->
      <tagsDecl>
        <rendition xml:id="italic">font-style:italic;</rendition>
        <namespace name="http://www.tei-c.org/ns/1.0">
          <tagUsage gi="hi" rendition="#italic"/>
        </namespace>
      </tagsDecl>
      <!-- ... -->
    </encodingDesc>
    <!-- ... -->
```

```
  </teiHeader>
  <text>
    <body>
      <!-- ... -->
      <p>'Plenty of options', he thought, sat on a bench and opened the book he had
       taken from the Whoopledywhaaian National Library. It was titled <hi>While thou
       art here</hi>, by Sir Edmund Peckwood. While reading the first sentence, his
       placid expression turned to a certain <hi>je ne sais quoi</hi>: 'For the first
       time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the
       annual meeting of the <hi rend="roman">Royal Academy of Whoopledywhaa</hi>.'</p>
      <!-- ... -->
    </body>
  </text>
</TEI>
```

Example 48. Referring to central definitions of rendition styles with `@rendition`.

5.9 `@cert`

The `@cert` attribute provides a method of indicating the encoder's certainty concerning an intervention or interpretation represented by the markup. This can be done with an informal classification, such as `"high"`, `"medium"`, or `"low"`, or more formal systems, such as a probability scale between `"1"` and `"0"`.

```
<choice xmlns="http://www.tei-c.org/ns/1.0">
  <sic cert="low">Pekwood</sic>
  <sic cert="high">Petwood</sic>
  <corr>Peckwood</corr>
</choice>
```

Example 49. Expressing certainty for an editorial interpretation with `@resp`.

In this example, two alternatives are presented for the transcription of the original form, with an indication of the certainty in their respective `@cert` attributes.

5.10 @resp

The @resp attribute is used to indicate the person or agency considered responsible for some aspects of the information encoded by an element. This responsible party should be identified formally in an element with an @xml:id attribute, either in the same document, or elsewhere.

```
<choice xmlns="http://www.tei-c.org/ns/1.0">
  <sic cert="low">Pekwood</sic>
  <sic cert="high">Petwood</sic>
  <corr resp="#EV">Peckwood</corr>
</choice>
```

Example 50. Identifying the person responsible for a correction with @resp.

5.11 @source

The @source attribute is used to indicate the source of an element and its content, for example by pointing to a bibliograhpic citation.

```
<text xmlns="http://www.tei-c.org/ns/1.0">
  <body>
    <!-- ... -->
    <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 <foreign>je ne sais quoi</foreign>:
    <quote source="#peckwood1935">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>
    <!-- ... -->
  </body>
  <back>
    <div type="bibliography">
      <listBibl>
        <!-- ... -->
```

```
        <bibl xml:id="peckwood1935"><author>Peckwood,
         Edmund</author>. <date when="1935">1935</date>. <title level="m">While thou art
         here</title>. <pubPlace>Whoopledywhaa</pubPlace>: <publisher>Whoopledywhaaian
         Press</publisher>.</bibl>
        <!-- ... -->
      </listBibl>
    </div>
  </back>
</text>
```

Example 51. Formally indicating the source for a quotation with `@source`.