TEI By Example

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 modified on: 2018-04-28
Licensed under a Creative Commons Attribution ShareAlike 3.0 License
TABLE OF CONTENTS

5. Global Attributes........................................................................................................................................... 1
  5.1 @xml:id .................................................................................................................................................... 1
  5.2 @n ............................................................................................................................................................ 2
  5.3 @xml:lang ................................................................................................................................................ 3
  5.4 @rend ....................................................................................................................................................... 3
  5.5 @rendition ............................................................................................................................................... 4
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. There are two kinds of attributes in the TEI world. Specific attributes are defined for specific elements only. They are documented in Appendix C. Elements of the TEI Guidelines. Global attributes, however, are optional attributes which are defined for every TEI element.

There are six global attributes:

- **@xml:id**: provides a unique identifier for an element.
- **@n**: provides a number or other label for an element. The number or label 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.
- **@rend**: indicates how the element was rendered or presented in the source text.
- **@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.

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 unique identifier as a value of 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:_TBEeg_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:_TBEeg_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:_TBEeg_id="p:2">He hoped his fellow members of the Royal Academy weren't even going to notice his absence.</p>`
2. 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 Whoopedywhaa.

3. 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 Royal Academy weren't even going to notice his absence.

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 illegal.

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:

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
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 only specify the language of an element when necessary for a specific purpose and different from the base language.

```xml
<p 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>
```

5.4 @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:

```xml
<p 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 <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>
```

......

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).

Both @xml:id and @n are used to express a reference system. This system can follow the text's canonical system, it may be derived from the structure of the electronic text, or it may be designed by the encoder. See Section 3.10.2 Creating New Reference Systems in the TEI Guidelines for a discussion of the latter.

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.
5.5 @rendition

Whereas the @rend attribute 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 teiHeader <teiHeader>, more specifically inside a <tagsDecl> inside <encodingDesc>. This is done in free text or using a formal language inside <rendition>. This way, only one description of the rendering must be given to which the @rendition attributes refer. 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> in the <tagsDecl> inside <teiHeader>, and a documentation of the deviation of that overall rendering in the third occurrence of <hi> in the text. 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 @render attribute of <tagUsage> refers to <rendition> by way of the latter's @xml:id attribute. This way, all <hi> elements inside <text> have 'italic' as their default style. The deviation to that rule is articulated in the third occurrence of <hi> by making use of the @rend attribute.

```xml
<teiHeader> <!-- ... --> <encodingDesc>
<tagsDecl>
  <rendition xml:id="IT">Italic</rendition>
  <namespace name="http://www.tei-c.org/ns/1.0">
    <tagUsage gi="hi" render="#IT"/>
  </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>
```