TEI Encoding patterns for citations, quotations and allusions: What's "in stock"?

International workshop on Computer Aided Processing of Intertextuality in Ancient Languages, June 2-4, 2014 - Lyon

Emmanuelle Morlock (UMR 5189 HISoMA)
emmanuelle.morlock@mom.fr
Model of text as an abstraction: 
Ordered 
Hierarchy of 
Content 
Objects

Customization explained formally and not formally in the ODD (one document does it all) = schema + documentation
WHERE does the TEI Bible Guidelines talk about citation?

TEI: P5 Guidelines

1. The TEI Infrastructure
2. The TEI Header
3. Elements Available in All TEI Documents
   3.1 Paragraphs
   3.2 Treatment of Punctuation
      3.2.1 Functions of Punctuation
      3.2.2 Hyphenation
   3.3 Highlighting and Quotation
      3.3.1 What is Highlighting?
      3.3.2 Emphasis, Foreign Words, and Unusual Language
         3.3.2.1 Foreign Words or Expressions
         3.3.2.2 Emphatic Words and Phrases
         3.3.2.3 Other Linguistically Distinct Material
      3.3.3 Quotation
      3.3.4 Terms, Glosses, Equivalents, and Descriptions
      3.3.5 Some Further Examples

TEI-L mailing list

<table>
<thead>
<tr>
<th>Item #</th>
<th>Date</th>
<th>Time</th>
<th>Lines</th>
<th>Subject</th>
<th>List Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>015799</td>
<td>2011-02-27</td>
<td>23:38</td>
<td>66</td>
<td>Re: Rendering and and friends</td>
<td>TEI-L</td>
</tr>
<tr>
<td>015801</td>
<td>2011-02-27</td>
<td>21:49</td>
<td>79</td>
<td>Re: Rendering and and friends</td>
<td>TEI-L</td>
</tr>
<tr>
<td>015800</td>
<td>2011-02-27</td>
<td>16:29</td>
<td>91</td>
<td>Re: Rendering and and friends</td>
<td>TEI-L</td>
</tr>
<tr>
<td>010082</td>
<td>2007-05-28</td>
<td>08:20</td>
<td>32</td>
<td>Re: q'nh (was TEI-Lite PS Problems with Scheme)</td>
<td>TEI-L</td>
</tr>
<tr>
<td>009247</td>
<td>2006-11-15</td>
<td>10:21</td>
<td>85</td>
<td>Re: q'nh (was TEI-Lite PS Problems with Scheme)</td>
<td>TEI-L</td>
</tr>
<tr>
<td>009249</td>
<td>2006-11-15</td>
<td>09:57</td>
<td>94</td>
<td>Re: q'nh (was TEI-Lite PS Problems with Scheme)</td>
<td>TEI-L</td>
</tr>
<tr>
<td>006473</td>
<td>2004-06-13</td>
<td>11:21</td>
<td>100</td>
<td>Re: Why so many quotation elements?</td>
<td>TEI-L</td>
</tr>
<tr>
<td>000260</td>
<td>1991-01-14</td>
<td>21:06</td>
<td>739</td>
<td></td>
<td>TEI-L</td>
</tr>
</tbody>
</table>

Feature requests and bug reports

Biblical references and quotations
### 3.3.3 quotation

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>q (quoted)</td>
<td>material which is distinguished from the surrounding text</td>
</tr>
<tr>
<td>said (speech or thought)</td>
<td>indicates passages thought or spoken aloud</td>
</tr>
<tr>
<td>quote (quotation)</td>
<td>contains a phrase or passage attributed by the narrator or author to some agency external to the text.</td>
</tr>
<tr>
<td>cit (cited quotation)</td>
<td>contains a quotation from some other document, together with a bibliographic reference to its source.</td>
</tr>
<tr>
<td>mentioned</td>
<td>marks words or phrases mentioned, not used.</td>
</tr>
<tr>
<td>soCalled</td>
<td>= scare quotes (distanciation)</td>
</tr>
</tbody>
</table>
16.2.3 Canonical references

```xml
<refsDecl>
  <cRefPattern matchPattern="<!-- regular expression -->"/>
</refsDecl>

<!-- ... -->

<p>The example is taken from <ref cRef="Gen. 1:1"/>.</p>
```
The TEI distinctions

- quotation introduced from the “quotation marks” perspective

- “quotation” vs “citation”

- `<quote>` as a specialisation of `<q>`
  => flexibility of the TEI to let you choose the level of precision you want...
**WHY would we want use it?**

**Point of markup:**
- make the pertinent features explicit (so it can be processed by a machine)
- add value through annotation
- facilitate sharing and re-use

**TEI Markup is a scholarly activity:**
- it’s never neutral (involves interpretation in all aspects)
- deciding what to markup and how (to support research questions) may be a research activity in itself
WHAT do we need to make explicit with citations?
WHAT do we need to make explicit with citations?
1. Explicit citation of Rm 6:8

“The Apostle said in his *Letter to the Romans* that we live and we die with Christ”
‘Citation’ in the broad sense may include:

<table>
<thead>
<tr>
<th></th>
<th>reproduced text</th>
<th>explicit reference</th>
<th>explicit attribution to a textual source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Explicit citations</td>
<td>++</td>
<td>++</td>
<td>+-</td>
</tr>
<tr>
<td>2. quotations with no attribution</td>
<td>++</td>
<td>Ø</td>
<td>Ø</td>
</tr>
<tr>
<td>2. Paraphrases</td>
<td>-</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>3. Allusions</td>
<td>-</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>4. Simple mentions (references strings)</td>
<td>Ø</td>
<td>+</td>
<td>-+</td>
</tr>
</tbody>
</table>
HOW - 1. explicit citation

- general mechanism:
  - `<quote>` demarcates the quotation
  - `<bibl>` or `<ref>` for the reference to the source
  - `<cit>` to enclose contiguous `<quote>` and `<bibl>` or `<ref>`
  - `pointers` (‘#ref01’) to associate non contiguous ones
How?
HOW - 1. explicit citation - example with <cit>

<cit>
  <ref cRef="Jon 2:1.b:heb">Heb :</ref>
  <quote>
    Et erat Ionas in uentre piscis tribus diebus et tribus noctibus.
  </quote>
</cit>
The Apostle said in his Letter to the Romans that we live and we die with Christ.
said> (...

<quote source="#Mc8.33 #Mt16.23 #Vg"> Vade, (...) retro Satana, scandalum es nobis. </quote>

(...)

</said>
HOW - 3. paraphrases

<seg type="paraphrase" source="#Gen 4:7">
  Peccasti, quiesce
</seg>

vs

<quote type="paraphrase" source="#Gen 4:7">
  Peccasti, quiesce
</quote>
HOW - 4. allusions

- allusion vs reminiscence
- implies blurry boundaries and difficulties in the analysis of the phenomenon:
  - textual referent? (whole books, episodes and not text, discontinuous elements...)
  - no explicit reference by definition
  - `<seg>` with @type? `<seg>` can have @source attribute...

`<seg type="allusion">`When jesus healed the blind man`</seg>`
HOW - 5. reference

- named entities, especially for characters and persons:
  - `<rs type="person">`<pensName ref="#pn01">David</persName>` and if the name means “book of Psalm”: `<seg type="reference" source="#psalm">`<persName ref="#pn01">David</persName>`</seg>`

- but can be useful for concepts or customary vocabulary
  - `<rs type="event">`he ascended to heaven`</rs>`
discontinuous elements: joins and standoff markup

- `<app>` to describe the comparison between the quotation and its textual referent
- `<certainty>`: formalized way of indicating degrees of precisions (or uncertainties)
- `@defective`: attribute of quote to indicate partial quotations (e.g. quotation of a part of a verse only)
Vade, inquit Sapientia, retro Satana, scandalum es nobis.

(said) (...)

<quote source="#Mc8.33 #Mt16.23">
Vade, inquit Sapientia, retro Satana, scandalum es nobis.
</quote>

(...)

</said>
HOW - Just for fun - 2) solution 1 - first step

(...)

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

(...)

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml:id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>

<quote xml:id="q1" source="#Mc8.33 #Mt16.23" next="#q2">Vade</quote>, inquit Sapientia,
<quote xml id="q2" source="#Mc8.33 #Mt16.23" prev="#q1">retro Satana, scandalum es nobis.</quote>
Just for fun: 2) solution 1 - step 2

(...)  
<quote xml:id="q1" source="#Mc8.33 #Mt16.23 #Vg">Vade</quote>,  
inquit Sapientia,  
<quote xml:id="q2" source="#Mc8.33 #Mt16.23 #Vg">retro Satana, scandalum es nobis.</quote>  
(...)

<!-- -->

<join target="#q1 #q2" result="quote" />
Just for fun : prefered solution

(...)

<quote xml:id="q1" source="#Mc8.33 #Mt16.23 #Vg">
Vade,
<seg type="incise" inquit Sapientia,></seg>
retro Satana, scandalum es nobis.
</quote>

(...)

<quote xml:id="q1" source="#Mc8.33 #Mt16.23 #Vg">
Vade,
<seg type="incise" inquit Sapientia,></seg>
retro Satana, scandalum es nobis.
</quote>

(...)

<quote xml:id="q1" source="#Mc8.33 #Mt16.23 #Vg">
Vade,
<seg type="incise" inquit Sapientia,></seg>
retro Satana, scandalum es nobis.
</quote>
Thanks !