
Treasures of the Taylorian - Documentation

Table of Contents

1. Editorial practice.....	1
2. Encoding practice.....	1

1. Editorial practice

Basic rules for encoding: The aim of the project is to publish short transcriptions that can easily be searched and the content is more important than the layout of the original. To give every student the chance to work on the aspects they think are most the interesting about their texts, some of the following rules give ample scope.

- Different forms of the letters *s*, *r*, *z* (long-s, long-z, r-rotunda etc.) are transcribed in the normal Latin form.
- Ligatures are solved (e.g. *æ* is transcribed as *ae*, *&* is transcribed as *et*).
- Overwritten letters are solved (*ŭ* is transcribed as *uo*).
- The letters *u/v* and *i/j* may either be preserved or normalised.
- Common abbreviations are usually solved without marking (e.g. nasal bars are solved as *n* or *m*, *XPI* is solved as *Christi* etc.). If the abbreviation is more complex or if the correct solution is not clear, the `<ex>` tag may be used.
- Layout specifics (a change in font size, indentation, capitalisation, underlining etc.) are not usually transcribed, but there may be exceptions.
- Capital letters and punctuation may either be preserved or normalised.
- Marginalia may be included in the transcription.
- Line breaks are not transcribed.

2. Encoding practice

The following is a list of all tags used in the transcriptions. They follow the TEI P5 guidelines and have been chosen to meet the project's specific needs. View the project's RelaxNG schema for more details.

Division `<div>`

Marks any structural unit (bigger than a paragraph), for example titlepage and rest-of-

the-book or single chapters.

@type="titlepage" is not mandatory, but can be used to define the division (e.g. titlepage, chapter, text).

Title <head>

Marks any heading or title in the text.

Paragraph <p>

Marks a paragraph (defined by content, not page layout, which means, a paragraph may include one or more page breaks).

Page break <pb/>

Marks the beginning of a new page.

@facs="001.jpg" (mandatory) - using the image file's name

@n="a1r" (mandatory) - using foliation, including r/v for recto/verso

Images <figure>

Marks a graphic, image, woodcut etc. in the original. The child tag <figureDesc> is used to describe the image and a title or heading can be given in <head>.

Closer <closer>

Marks the closing words that sometimes can be found at the end of early prints or letters. It may contain names, dedications, salutes, dates or places. Usually, the closer is highlighted in the layout by indent, centring or some such.

Imprint <imprint>

Marks an imprint and usually includes the name of a printer (<publisher><rs type="person" ref="#...">...</rs></publisher>), a place (<pubPlace>) and a date (<date>).

Date <date>

Marks any date anywhere in the text or the metadata.

@when="20151029" (mandatory) displays the date in a machine readable form (datatype: „integer“). Usually, if only the year is known, you put two zeros for both the month and the day (same if only the year and month are known, so 20151000 for „Oct. 2015“).

Marginalia <note place=“margin“>

Marks marginalia in the original, either printed or added by hand. If necessary, more details can be added to describe the exact location of the marginalia on the page, the scribe, the type of information (summary, bibliographical information etc.) and many more.

@place="margin" is mandatory.

Footnotes <note type=“footnote“>

Marks a comment by the modern editor, not something in the original, e.g. a comment on an unreadable part of the text or basically just any critical comment.

@type="footnote" is mandatory.

Abbreviations <ex>

Marks an abbreviation that has been solved by the editor. E.g. *ē* could be solved as *est*. The <ex> marks the letters added by the editor: e<ex>st</ex>

Text decoration / layout <hi>

Marks a highlighted part of the text in the original. The type of highlighting must be defined in **@rend**, e.g. **@rend**="underline / initial / smallcaps /..."

Line breaks <lb/>

Usually, line breaks are not encoded. But if necessary (e.g. in a heading), the <lb/> tag stands at the beginning of each new line.

Poems <lg>+<l>

Marks poems as a whole, whereas inside the <lg> element the <l> element is used to mark the individual lines. If <lg> is used the <l> elements are compulsory.

Person <rs type="person">

Marks a person in the **original text**. The @type="person" is mandatory and so is the @ref, which is used to refer to the list of persons (<listPerson>) included in the <teiHeader>. For example <rs type="person" ref="#luther_martin">Lutther</rs> in the transcription refers to the entry <person xml:id="luther_martin">...</person> in the <teiHeader>.

NB: The following three tags are used to encode people: <rs type="person" ref="#...">, <person> and <persName>. The main difference is that the first one is used to mark people's names in the original text, the other two are for metatexts only:

<rs type="person" ref="#..."> is used to mark people's names in the original text with the @ref referring to a normalised entry in the <listPerson>.

<person> may contain many different types of information about a person, such as name, address, email etc. It's only used in <listPerson> in the <teiHeader>.

<persName> contains only a name (e.g. of an author, editor or transcriber) and it's only used in the <teiHeader>, never the original text.

Forme Work <fw>

Marks custodes, foliation etc. in the original text.

List <list>

Marks a list of items (<item>), eg. theses. Numbers or labels may be added in <label>.

Last modified by Jennifer Bunselmeier, 15.01.2016