

MARKUP INSTRUCTIONS FOR CONTRIBUTORS

(ABBOTT-SMITH LEXICON)

2012-09-27

This file describes and illustrates the TEI markup practices employed in this project. To contribute you must be familiar with the basics of how XML works. For an introduction, see http://www.w3schools.com/xml/xml_what.asp.

This project follows several markup practices. Here are several of the most important ones:

I. Structure of Lexicon Entries

- A. Every entry is nested in an <entry> element with an id (@n), which should be the Greek lemma. If there are multiple options at the head of the entry, choose the first.
- B. Nested within <entry> are several elements: <form>, <gramGrp>, and <sense>. The first two only appear once. The last one may appear multiple times if there are multiple sense numbers. Note that all text in the entry must fall within one of these three elements, not directly under <entry>.

II. Specific Elements of Each Entry

A. <form>

- i. Typically this will include all information prior to the first definition.
- ii. A comma usually separates the form information from sense information.

B. <gramGrp>

- i. This element is for grammatical information that does not involve the form of the word.
- ii. This normally will have nothing in it, in which case it can be dropped or included in this way: <gramGrp/>.
- iii. If some grammatical information is given in the entry such as part of speech <pos> or some other subcategorization <subc>, it would be marked up in this way: <gramGrp><pos>verb</pos>, <subc>intransitive</subc></gramGrp>

C. <sense>

- i. Anything related to the meaning of a word should occur in this element.

- ii. Glosses (usually anything in italics) should also be within <gloss>.
- iii. If sense numbers are given, include it in @n in <sense> but not in the text as it is written in the lexicon. That can be included later using XSL.
- iv. Ideally all <sense> elements should be nested within a single <sense> element.

III. Other Elements

A. <pb />

- i. All page breaks are included in the file.
- ii. After cleaning up a page, add your initials to the commented @typed field after the <pb> element. After checking a page, do the same under @checked. This is not valid TEI, so the attributes are commented out outside the <pb> element.
- iii. A completed page should look like this: <pb n="1" /> <!-- typed="ABC" checked="XYZ" -->

B. <foreign>

- i. All text is assumed to be English (the analysis language) unless otherwise specified.
- ii. The <foreign> element can be used for this purpose.
- iii. Examples include:
 - a) Greek: <foreign xml:lang="grc">'Ααρών</foreign>
 - b) Aramaic: <foreign xml:lang="arc">ܐܪܡܝܐ</foreign>
 - c) Hebrew: <foreign xml:lang="heb">יְהוָה</foreign>
 - d) Latin: <foreign xml:lang="lat">diligo</foreign>
- iv. If you do not know how to type a given language, such as Hebrew, include the following: <foreign xml:lang="heb">Hebrew</foreign>. Someone else can add the Hebrew later.

C. <ref>

- i. Biblical references should be marked up using this element.

- ii. Although Abbott-Smith uses superscript numbers for verses, this project does not. Instead, we use a colon to separate chapter and verse.
- iii. For biblical book names in @osisRef, use SBL abbreviations. See <http://www.textonline.org/textstyleguidelines>.
- iv. Examples include:
 - a) <ref osisRef="Luke.1.5">Lk 1:5</ref>
 - b) <ref osisRef="Mark.1.14-Mark.1.15">Mk 1:14-15</ref>
 - c) <ref osisRef="1Chr.24.3 1Chr.24.10">I Ch 24:3, 10</ref>

IV. Recurring Text

- A. For a dagger use † (see Preface of the lexicon for the meaning of this symbol)
- B. For < use <
- C. For > use >
- D. For superscript, use: <hi rend="superscript">
- E. For subscript, use: <hi rend="subscript">
- F. For an ampersand use &
- G. For paragraphs use <p>, and for line breaks within elements use <lb />
- H. For italics that are not for a definition of the current entry, use <emph>. This includes abbreviations of published sources as well as definitions of related terms. For examples, see <entry n="Ἀβαδδών"> and <entry n="ἁγάπη">.

EXAMPLE MARKUP

<entry n="α">

<form><foreign xml:lang="grc">Α, α, ἄλφα</foreign> (q.v.), <foreign xml:lang="grc">τό</foreign>, indecl., </form>

<gramGrp/>

<sense><gloss>alpha</gloss>, the first letter of the Greek alphabet. As a numeral, <foreign xml:lang="grc">ἄ</foreign> = 1, <foreign xml:lang="grc">α</foreign> = 1000. As a prefix, it appears to have at least two and perhaps three distinct senses:

<sense n="1">

<foreign xml:lang="grc">ἄ</foreign> (before a vowel, <foreign xml:lang="grc">ἄν</foreign>) <gloss>negative</gloss>, as in <foreign xml:lang="grc">ἄγνωστος, ἄδικος</foreign>.

</sense>

<sense n="2">

<foreign xml:lang="grc">ἄ-, ἄ</foreign> <gloss>copulative</gloss>, indicating community and fellowship, as in <foreign xml:lang="grc">ἄ-πλοῦς, ἄ-κολουθέω, ἄ-δελφός</foreign>.

</sense>

<sense n="3">An intensive force (LS, s. <foreign xml:lang="grc">α</foreign>), as in <foreign xml:lang="grc">ἄ-τενίζω</foreign> is sometimes assumed (but v. Boisacq, s.v.).</sense>

</sense>

</entry>

<entry n="Ἀαρόν">

<form><foreign xml:lang="grc">Ἀαρόν</foreign> (Heb. <foreign xml:lang="heb">אַהֲרֹן</foreign>), indecl. (in FLJ, <foreign xml:lang="grc">-ῶνος</foreign>), </form>

<gramGrp/>

<sense><gloss>Aaron</gloss> (<ref osisRef="Exod.4.14">Ex 4:14</ref>, al.): <ref osisRef="Luke.1.5">Lk 1:5</ref>, <ref osisRef="Acts.7.40">Ac 7:40</ref>, <ref osisRef="Heb.5.4">He 5:4</ref>, <ref osisRef="Heb.7.11">7:11</ref>, <ref osisRef="Heb.9.4">9:4</ref>.†

</sense>

</entry>

ENTRY TEMPLATE

```

<entry n="">

    <form><foreign xml:lang="grc"></foreign></form>

    <gramGrp></gramGrp>

    <sense>

        <gloss></gloss>

        <refosisRef=""></ref>

    </sense>

    <sense n="1">

        †

    </sense>

</entry>

```