

Presentation structure:

2) Outline of our schema

1st Level: super entry / entry

1.1 Super entry:

Head word (ex. verb and compounds).

1.2 Entry:

Homographs are in different entries. (according to TEI they can also be put in the same entry ^ with <hom n="1">)

Attributes:

2nd Level: form, gramGrp, etym, xr, sense, note

2.1 Form:

Form is a property of “entry”. It is not allowed for superentry (it is in TEI)

Multiple forms are allowed (dialects, inflections - examples with attributes).

2.1.1. Usg: dialectal information, refers to form only. According to TEI , “usg” can be associated with different senses:

2.1.2 Orth: multiple “orth” possible

2.1.4 GramGrp

Same form, different pos – different entries. (can be in the same entry according to TEI)

2.2 Gramgrp

GramGrp can refer to either entry or form, it is not allowed on “sense” level (this is allowed by TEI). Only one “GramGrp” is allowed per form or per entry. But we do allow values like “m./f.”; theoretically we could encode them in separate GramGrps.

2.2.1 pos

2.2.3 subc

2.2.4 gen

2.2.5 gram

2.2.6 number

2.3 Etym : the structure does not correspond to TEI. What is the difference between “ref” and “xr” here?. “XR” also contains “ref”

2.3.1 ref

2.3.2 xr

2.4 xr

2.4.1 ref: mention that it will be updated to refer to a unique ID. Also mention that “xr” is a more reliable way to refer to a head word and compounds constituents. Maybe it makes sense to add “lbl” to “xr”

2.5. Sense

multiple senses are allowed.

“sense” does not allow grammatical information

2.5.1 ref

2.5.2 cit: attribute type=translation, and lang

2.5.2.1 quote

2.5.2.2 def: does not need an attribute type=translation in “cit”

2.5.2.3 bibl