1) Homographs:

- Some dictionaries provide distinct entries for homographs, on the basis of etymology, part-of-speech, or both, and typically provide a numeric superscript on the headword identifying the homograph number. In these cases each homograph should be encoded as a separate entry;

- the superEntry element may optionally be used to group such successive homograph entries
- “form” can appear at the superentry level

Distinct entry for homograph

```
<superEntry>
  <entry type="hom" n="1">
   <form>
    <orth>grève</orth>
<!-- strike -->
   </form>
  </entry>
  <entry type="hom" n="2">
   <form>
    <orth>grève</orth>
<!-- shore -->
   </form>
  </entry>
 </superEntry>
```

“form” can appear at the superentry level

```
<superEntry>
 <form>
  <orth>abandon</orth>
  <hyph>a|ban|don</hyph>
  <pron>@"band@n</pron>
 </form>
 <entry n="1">
  <gramGrp>
   <pos>v</pos>
   <subc>T1</subc>
  </gramGrp>
  <sense n="1">
   <def>to leave completely and for ever … </def>
  </sense>
  <sense n="2"/>
 </entry>
 <entry n="2">
  <gramGrp>
   <pos>n</pos>
   <subc>U</subc>
  </gramGrp>
  <sense>
   <def>the state when one's feelings and actions are
uncontrolled; freedom
      from control…</def>
  </sense>
 </entry>
</superEntry>
```

- If homographs are different parts of speech (a walk, to walk):

The hom element marks the subdivision of entries into homographs differing in their part-of-speech. The sense element marks the subdivision of entries and part-of-speech homographs into senses; (with sub-senses):

<entry> <hom n="1"> <sense n="1"> <!-- ... --> </sense> <sense n="2"> <!-- ... --> </sense> </hom> <hom n="2"> <sense n="1"> <sense n="a"> <!-- ... --> </sense>	<entry> <form> <orth>bray</orth> <pron>breI</pron> </form> <hom> <gramGrp> <pos>n</pos> </gramGrp> <sense> <def>cry of an ass; sound of a trumpet.</def> </sense> </hom> <hom> <gramGrp> <pos>vt</pos> <subc>VP2A</subc> </gramGrp> <sense> <def>make a cry or sound of this kind.</def> </sense> </hom> </entry>

2) According to TEI not only "entry" or "hom" can constitute a top level, but also a "sense":

Any of the hierarchical levels (entry, entryFree, hom, and sense) may contain any of these top-level constituents[1], since information about word form, particular grammatical information, special

- [1] form (form information group) groups all the information on the written and spoken forms of one headword.
- gramGrp (grammatical information group) groups morpho-syntactic information about a lexical item, e.g. pos, gen, number, case, or iType (inflectional class).
- def (definition) contains definition text in a dictionary entry.
- cit (cited quotation) contains a quotation from some other document, together with a bibliographic reference to its source. In a dictionary it may contain an example text with at least one occurrence of the word form, used in the sense being described, or a translation of the headword, or an example.
- usg (usage) contains usage information in a dictionary entry.
- xr (cross-reference phrase) contains a phrase, sentence, or icon referring the reader to some other location in this or another text.
- etym (etymology) encloses the etymological information in a dictionary entry.
- re (related entry) contains a dictionary entry for a lexical item related to the headword, such as a compound phrase or derived form, embedded inside a larger entry.
- note contains a note or annotation.

pronunciation, usage information, etc., may apply to an entire entry, or to only one homograph, or only to a particular sense.

In the following example “sense” contains not only a different “def”, but also an additional grammatical subcategory:

```
<entry>
 <form>
  <orth>careen</orth>
  <hyph>ca|reen</hyph>
  <pron>k@"ri:n</pron>
 </form>
 <gramGrp>
  <pos>vt</pos>
  <pos>vi</pos>
 </gramGrp>
 <sense n="1">
  <gramGrp>
   <subc>VP6A</subc>
  </gramGrp>
  <def>turn (a ship) on one side for cleaning,
repairing, etc.</def>
 </sense>
 <sense n="2">
  <gramGrp>
   <subc>VP6A</subc>
   <subc>VP2A</subc>
  </gramGrp>
  <def>(cause to) tilt, lean over to one side.</def>
 </sense>
</entry>
```

3) “GramGrp” may occur within “form” with corresponding grammatical information[2]

[2] • gen (gender) identifies the morphological gender of a lexical item, as given in the dictionary.

- number indicates grammatical number associated with a form, as given in a dictionary.
- case contains grammatical case information given by a dictionary for a given form.
- per (person) contains an indication of the grammatical person (1st, 2nd, 3rd, etc.) associated with a given inflected form in a dictionary.
- tns (tense) indicates the grammatical tense associated with a given inflected form in a dictionary.
- mood contains information about the grammatical mood of verbs (e.g. indicative, subjunctive, imperative).

4) Multiple orthographic forms may be given, e.g. to illustrate a word's inflectional pattern:
- forms can be nested

<form> <orth>brag</orth> </form> <gramGrp> <pos>vb</pos> </gramGrp> <form type="inflected"> <orth>brags</orth> <orth>bragging</orth> <orth>bragged</orth> </form>	<form> <orth>hospitaller</orth> <form> <usg type="geo">U.S.</usg> <orth>hospitaler</orth> </form> <pron>"hQspIt@l@</pron> </form>

5) The gramGrp element groups grammatical information ... **and can appear as a child of entry, form, sense, cit, or any other element** containing content about which there is grammatical information.

Coding with "gram" possible <form> <orth>médire</orth> </form> <gramGrp> <gram type="pos">v</gram> <gram type="subc">t ind</gram> <gram type="collocPrep">de</gram> </gramGrp>	**Multiple "gramGrp" for entry:** <form> <orth>isotope</orth> </form> <gramGrp> <pos>adj</pos> </gramGrp> <gramGrp> <pos>n</pos> <gen>m</gen> </gramGrp>
"GramGrp" in "sense": <entry> <form>	

- iType (inflectional class) indicates the inflectional class associated with a lexical item.

type indicates the type of indicator used to specify the inflection class, when it is necessary to distinguish between the usual abbreviated indications (e.g. inv) and other kinds of indicators, such as special codes referring to conjugation patterns, etc. Beispielwerte sind etwa: 1] abbrev; 2] verbTable

- pos (part of speech) indicates the part of speech assigned to a dictionary headword such as noun, verb, or adjective.
- subc (subcategorization) contains subcategorization information (transitive/intransitive, countable/non-countable, etc.)
- colloc (collocate) contains any sequence of words that co-occur with the headword with significant frequency.

<orth>wits</orth> <pron>wIts</pron> </form> <gramGrp> <number>pl</number> <pos>n</pos> </gramGrp> <sense n="1"> <gramGrp> <number>sometimes sing.</number> </gramGrp> <def>the ability to reason and act, esp. quickly …</def> </sense> </entry>	

6) Definitions may occur directly within an entry; when multiple definitions are given, they are typically identified as belonging to distinct senses, as here:

<entry> <form> <orth>demigod</orth> <pron> … </pron> </form> <gramGrp> <pos>n</pos> </gramGrp> <sense n="1"> <sense n="a"> <def>a being who is part mortal, part god.</def> </sense> <sense n="b"> <def>a lesser deity.</def> </sense> </sense> <sense n="2"> <def>a godlike person.</def> </sense> </entry>	but can be also inside of “cit”: <entry> <form> <orth>rémoulade</orth> <pron>Remulad</pron> </form> <gramGrp> <pos>n</pos> <gen>f</gen> </gramGrp> **<cit type="translation" xml:lang="en">** **<quote>remoulade</quote>** **<quote>rémoulade</quote>** **<def>dressing containing mustard and herbs</def>** </cit> </entry>

7) “Usg” can be used with “sense”:

In the next example, usage labels are used to indicate domains, register, and synonyms associated with different senses:

<sense n="a"> <usg type="dom">Peinture</usg> <usg type="style">lit</usg> <usg type="style">fig</usg> <cit type="translation" xml:lang="en"> <quote>palette</quote> </cit> </sense> <sense n="b"> <usg type="dom">Boucherie</usg> <cit type="translation" xml:lang="en"> <quote>shoulder</quote> </cit> </sense> <sense n="c"> <sense> <usg type="syn">aube de roue</usg> <cit type="translation" xml:lang="en"> <quote>paddle</quote> </cit> </sense> <sense> <usg type="syn">battoir à linge</usg> <cit type="translation" xml:lang="en"> <quote>beetle</quote> </cit> </sense> <sense> <usg type="dom">Manutention</usg> <usg type="dom">Constr</usg> <cit type="translation" xml:lang="en"> <quote>pallet</quote> </cit> </sense> </sense>	

8) Possibility for reference

<oRef> (orthographic-form reference) in a dictionary example, indicates a reference to the orthographic form(s) of the headword.

<pRef> (pronunciation reference) in a dictionary example, indicates a reference to the pronunciation(s) of the headword.

<oVar> (orthographic-variant reference) in a dictionary example, indicates a reference to variant orthographic form(s) of the headword.

<entry> <form> <orth xml:id="di-o1">vag-</orth> <orth xml:id="di-o2">vago-</orth> </form> <def>vagus nerve</def> <cit type="example"> <quote> <oRef target="#di-o1" type="nohyph"/>al</quote> <quote> <oRef target="#di-o2" type="nohyph"/>tomy</quote> </cit> </entry>	