# Taggar till SOL

**Taggar att använda:**

<div> omger en post och innehåller typerna ”bebyggelsenamn”, ”naturnamn” eller ”namnelement”, beroende på om det är ett bebyggelsenamn eller ett naturnamn, och xml:id, som är ortnamnet om det är unikt, annars använder vi ”Ekeby1”, ”Ekeby2” osv.

<head> omger uppslagsnamnet

<p> omger ett stycke

<placeName> omger ortnamn, dels uppslagsnamnet, dels andra ortnamn som omnämns i texten, t.ex. Ryssby nedan

<locale> omger denotationen, dvs vad det är för lokal som bär namnet

<location> omger den geografiska information som anger var man finner namnet ifråga

<district> är en undergrupp inom <location> och innehåller typerna ”socken” och ”härad”

<region> är en undergrupp inom <location> och innehåller typen ”landskap”

<settlement> är en undergrupp inom <location> och innehåller typen ”city”

**Viktigt** <location> får inte innehålla text!

Det kan finnas tre olika typer av information i kursiv stil:

1. Belägg, som är uppmarkerade med <span type=’belägg’>
2. Etymon, markerade med <span type=’etymon’> eventuellt även har attributet ’lang’ med värden ”smi” for samiska, ”sma” för sydsamiska, ”sju” för umesamiska, ”sje” för pitesamiska, ”smj” för lulesamiska, ”se” för nordsamiska, ”fi” för finska och ”fit” för meänkieliska.
3. Ortnamn, markerade med <placeName> som ovan.

**Exempel**

<div type=”bebyggelsenamn” xml:id=”Rockneby”>

<head><placeName>**Rockneby**</placeName**>**</head> <p><span type=’locale’>tätort</span>, <location><district type=”socken”>Ryssby sn</district><district type=”härad”>Norra Möre hd</district><region type=”landskap”>Småland</region></location>. *Rocnaby* 1312. – Tätorten har vuxit fram kring en järnvägsstation vid en 1897 öppnad sträcka. Stationen, som fick sitt namn trots att den inte anlades på byn Rocknebys mark, har säkerligen övertagit namnet från den i denna by belägna poststationen, inrättad 1874. Att poststationen trots grannskapet till sockenkyrkan och kyrkbyn Ryssby inte namngavs efter socknen sammanhänger troligast med att <placeName>*Ryssby*</placeName>redan var reserverat för den år 1878 inrättade järnvägs- och poststationen i Ryssby socken i Sunnerbo hd, också den i Småland. Förleden i *Rockneby* innehåller det från Småland kända dialektordet *racken* ’låg, sammanhängande ås’. Efterleden är → *by* ’gård; by’. </p> </div>

<div type=”bebyggelsenamn” xml:id=”Barturtte”>

<head><placeName>**Barturtte**</placeName**>**</head> <p> <locale>fjäll</locale> <location>vid Kungsleden, <district type=”socken”>Arjeplogs sn</district>, <region type=”landskap”>Lappland</region>.</location> – Det pitesamiska namnets förled innehåller kanske ordet <lang type=”etymon” lang=’sje’>*bartta*</span>’(rök)stuga’. Efterledens betydelse är okänd. Samma förled ingår i namnet på fjällsjön <placeName>*Bartavrre*</placeName**>**intill fjället (tidigare skrivet <placeName>*Bartaure*</placeName**>**). Sjönamnets efterled är en förkortad form av pitesam. <lang type=”etymon” lang=”sje:>*javrre* ’sjö’</span>. </p> </div>

<div type=”namnelement” xml:id=”by”>

<head>**by**</head><p>Ortnamn på -*by* förekommer i alla landsändar. De är talrikast i Mälarlandskapen och Östergötland. En större koncentration finns också i området väster om Klarälven– Vänern–Göta älv. Deras sammanlagda antal i Sverige har uppskattats till ca 3 400. Grundordet *by* sammanhänger med verbet *bo*, kanske i dess äldre betydelse ’bereda, ställa i ordning’. *By* kan ursprungligen ha betytt ’beredning, anordning’ och mer konkret ’jordområde där man gjort anordningar (hus, röjning, odling) för att avvinna jorden nyttigheter i form av bete, gröda osv.’ En utbredd uppfattning har länge varit att *by* redan från början betecknat bebyggelser. Det är dock mycket möjligt att de allra äldsta, forntida ortnamnen på -*by* i stället avsett (ny)odlingar; enligt en mening har det varit fråga om lövängar. Ortnamnselementet bör hur som helst tidigt ha kommit att brukas om bebyggelser, och i flertalet fall har man att räkna med innebörden ’gård’ eller ’by’. Elementet *by* har använts om bebyggelser som tillkommit genom utvidgning av, utflyttning från eller delning av äldre moderbebyggelser. Huruvida man står inför den äldre betydelsen ’gård’ eller inför ’by, samling av gårdar’ får avgöras från fall till fall. Ortnamnstypen -*by* har varit produktiv under en mycket lång tid, från forntiden och in i nya tiden. Den bortre gränsen för *by*-namnens bildningstid förlades av en äldre forskning till folkvandringstiden (400–550 e.Kr.), men i senare tid har det hävdats att namntypens äldsta företrädare går tillbaka ända till romersk järnålder (Kr.f.–400 e.Kr.). </p>

<p>Förlederna i -*by*-namnen är ofta naturbeskrivande eller terrängbetecknande, så t.ex. i sockennamnen →<placeName>*Asby*</placeName>, →<placeName>*Dalby*</placeName>, →<placeName>*Ekeby*</placeName>, →<placeName>*Hammarby*</placeName>, →<placeName>*Hässleby*</placeName>och →<placeName>*Säby*</placeName>. Sockennamnet →<placeName>*Vimmerby*</placeName>innehåller sannolikt ett gammalt ånamn. Andra förleder anger storlek, relativ ålder eller läge; exempel härpå ger sockennamnet →<placeName>*Myckleby*</placeName>, det på olika håll uppträdande <placeName>*Nyby*(*n*) </placeName> och sockennamn som →<placeName>*Mellby*</placeName> →<placeName>*Norrby*</placeName>och →<placeName>*Sörby*</placeName>. På en helig plats syftar →<placeName>*Viby*</placeName>. Ett mångenstädes uppträdande namn är →<placeName>*Broby*</placeName>, som i några fall blivit sockennamn. I många fall är förleden ett personnamn eller en personbeteckning. Namn på -*by* innehållande personnamn har ansetts vara relativt unga; sådana förekommer bl.a. i Västsverige och Uppland. Gamla bildningar är dock sockennamnen →<placeName>*Karleby*</placeName>, →<placeName>*Rinkaby*</placeName>och →<placeName>*Tegneby*</placeName>, som antas innehålla beteckningar för män i forntida konungars tjänst.

<p>Personbetecknande är också den på olika håll förekommande förledstyp som består av ett mansnamn åtföljt av ordet fsv. *arve* ’arvtagare, son’ i genitiv (→ *arve*); ett exempel bland många är bynamnet <placeName>*Tosarby*</placeName>(*Thorisaruaby* 1384) ’Tores arvinges gård’ i Gustafs socken, Dalarna. En markerad mellansvensk grupp utgör de -*by*-namn som i förleden innehåller en inbyggarbeteckning sammansatt med -*karl*; ett bekant hithörande namn är →<placeName>*Älvkarleby*</placeName>’älvkarlarnas, älvbornas gård eller by’. </p>

<p>I vissa fall har -*by* i efterhand lagts till ett äldre ortnamn; jfr diskussionen om → *husaby*. </p>

<p>Litteratur: Lars Hellberg, Inbyggarnamn på -karlar i svenska ortnamn 1 (1950); dens., Kumlabygdens ortnamn och äldre bebyggelse (1967), s. 243–57, 369–419; Elof Hellquist, De svenska ortnamnen på by. En översikt (1918); Tom Schmidt, Norske gårdsnavn på -*by* og -*bo* med personnavnforledd 1–2 (2000). </p></div>

# Utdrag från TEI

#### 13.2.3 Place Names

Like other proper nouns or noun phrases used as names, place names can simply be marked up with the [rs](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-rs.html) element, or with the [name](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-name.html) element. For cartographers and historical geographers, however, the component parts of a place name provide important information about the relation between the name and some spot in space and time. They also provide important evidence in historical linguistics.

These Guidelines distinguish three ways of referring to places. A place name (represented using the [placeName](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-placeName.html) element) may consist of one or more names for hierarchically-organized geo-political or administrative units (see section [13.2.3.1 Geo-political Place Names](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html#NDPLGU)). A place named simply in terms of geographical features such as mountains or rivers is represented using the [geogName](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-geogName.html) element (see section [13.2.3.2 Geographic Names](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html#NDPLGF)). Finally, an expression consisting of phrases expressing spatial or other kinds of relationship between other kinds of named place may itself be regarded as a way of referring to a place, and hence as a kind of named place (see section [13.2.3.3 Relative Place Names](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html#NDPLR)).

* [placeName](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-placeName.html) contains an absolute or relative place name.
* [geogName](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-geogName.html) (geographical name) identifies a name associated with some geographical feature such as Windrush Valley or Mount Sinai.

As members of the [att.naming](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-att.naming.html) class, all of these elements bear the attributes key, ref, and nymRef mentioned above. These attributes are primarily useful as a means of linking a place name with information about a specific place. Recommendations for the encoding of information about a place, as distinct from its name, are provided in [13.3.4 Places](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html#NDGEOG) below.

Like the [persName](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-persName.html) element discussed in section [13.2.1 Personal Names](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html#NDPER), the [placeName](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-placeName.html) element may be regarded simply as an abbreviation for the elements <name type="place"> or <rs type="place">. The following encodings are thus equivalent:48

After  
spending some time in our <rs ref="tag:projectname.org,2012:NY1"  
 type="place">modern <name ref="tag:projectname.org,2012:BA1"  
  type="place">Babylon</name>  
</rs>, <name ref="tag:projectname.org,2012:NY1"  
 type="place">New York</name>, I have proceeded to  
the <rs ref="tag:projectname.org,2012:PH1"  
 type="place">City of Brotherly Love</rs>.

After spending some  
time in our <placeName ref="tag:projectname.org,2012:NY1">modern  
<placeName ref="tag:projectname.org,2012:BA1">Babylon</placeName>  
</placeName>,  
<placeName ref="tag:projectname.org,2012:NY1">New  
York</placeName>, I have proceeded to the <placeName ref="tag:projectname.org,2012:PH1">City of  
Brotherly Love</placeName>.

**13.2.3.1 Geo-political Place Names**

A place name may contain text with no indication of its internal structure:

<placeName>Rochester,  
NY</placeName>

More usually however, a place name of this kind will be further analysed in terms of its constitutive geo-political or administrative units. These may be arranged in ascending sequence according to their size or administrative importance, for example: ‘Rochester, New York’, or as a single such unit, for example ‘Belgium’. These Guidelines provide a hierarchy of generic element names, each of which may be more exactly specified by means of a type attribute:

* [district](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-district.html) contains the name of any kind of subdivision of a settlement, such as a parish, ward, or other administrative or geographic unit.
* [settlement](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-settlement.html) contains the name of a settlement such as a city, town, or village identified as a single geo-political or administrative unit.
* [region](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-region.html) contains the name of an administrative unit such as a state, province, or county, larger than a settlement, but smaller than a country.
* [country](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-country.html) contains the name of a geo-political unit, such as a nation, country, colony, or commonwealth, larger than or administratively superior to a region and smaller than a bloc.
* [bloc](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-bloc.html) contains the name of a geo-political unit consisting of two or more nation states or countries.

These elements are all members of the [model.placeNamePart](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-model.placeNamePart.html) class, members of which may be used anywhere that text is permitted, including within each other as in the following examples:

<placeName>  
 <settlement type="city">Rochester</settlement>,  
<region type="state">New York</region>  
</placeName>

<placeName ref="tag:projectname.org,2012:LSEA1">  
 <country type="nation">Laos</country>,  
<bloc type="sub-continent">Southeast Asia</bloc>  
</placeName>

<placeName>  
 <district type="arondissement">6ème</district>  
 <settlement type="city">Paris, </settlement>  
 <country>France</country>  
</placeName>

#### 13.3.4 Places

In [13.2.3 Place Names](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html#NDPLAC) we discuss various ways of naming places such as towns, countries, etc. In much the same way as these Guidelines distinguish between the encoding of names for people and the encoding of other data about people, so they also distinguish between the encoding of names for places and the encoding of other data about places. In this section we present elements which may be used to record in a structured way data about places of any kind which might be named or referenced within a text. Such data may be useful as a way of normalizing or standardizing references to particular places, as the raw material for a gazetteer or similar reference document associated with a particular text or set of texts, or in conjunction with any form of geographical information system.

The following elements are provided for this purpose:

* [listPlace](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-listPlace.html) (list of places) contains a list of places, optionally followed by a list of relationships (other than containment) defined amongst them.
* [place](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-place.html) contains data about a geographic location

The [model.placeStateLike](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-model.placeStateLike.html) class contains elements describing characteristics of a place which have a definite duration, such as its name. Any member of the [model.placeNamePart](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-model.placeNamePart.html) may be used for this purpose, since a [place](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-place.html) element will usually contain at least one, and possibly several, [placeName](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-placeName.html)-like elements indicating the names associated with it, by different people, in different languages, or at different times.

For example, the modern city of Lyon in France was in Roman times known as Lugdunum. Although the modern and the Roman city are not physically co-extensive, they have significant areas which overlap, and we may therefore wish to regard them as the same place, while supplying both names with an indication of the time period during which each was current.

A place is defined, however, by its physical location, which does not typically change over time. Locations may be specified in a number of ways: as a set of coordinates defining a point or an area on the surface of the earth, or by providing a description of how the place may be found, usually in terms of other place names. For example, we can identify the location of the Canadian city of London, either by specifying its latitude and longitude, or by specifying that we mean the city called London located in the province called Ontario within the country called Canada.

In addition we may wish to supply a brief characterization of the place identified, for example to state that it is a city, an administrative area such as a country, or a landmark of some kind such as a monument or a battlefield. If our typology of places is simple, the open ended type attribute is the easiest way to represent it: so we might say <place type="city">, <place type="battlefield"> etc.

Within the [place](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-place.html) element, the following elements may be used to provide more information about specific aspects of the place in a structured form:

* [placeName](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-placeName.html) contains an absolute or relative place name.
* [location](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-location.html) defines the location of a place as a set of geographical coordinates, in terms of other named geo-political entities, or as an address.

### 9.2 The Structure of Dictionary Entries

A simple dictionary entry may contain information about the form of the word treated, its grammatical characterization, its definition, synonyms, or translation equivalents, its etymology, cross-references to other entries, usage information, and examples. These we refer to as the constituent parts or constituents of the entry; some dictionary constituents possess no internal structure, while others are most naturally viewed as groups of smaller elements, which may be marked in their own right. In some styles of markup, tags will be applied only to the low-level items, leaving the constituent groups which contain them untagged. We distinguish the class of top-level constituents of dictionary entries, which can occur directly within the [entry](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-entry.html) element, from the class of phrase-level constituents, which can normally occur only within top-level constituents. The top-level constituents of dictionary entries are described in section [9.2.2 Groups and Constituents](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DIENGP), and documented more fully, together with their phrase-level sub-constituents, in section [9.3 Top-level Constituents of Entries](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITP).

In addition, however, dictionary entries often have a complex hierarchical structure. For example, an entry may consist of two or more sub-parts, each corresponding to information for a different part-of-speech homograph of the headword. The entry (or part-of-speech homographs, if the entry is split this way) may also consist of senses, each of which may in turn be composed of two or more sub-senses, etc. Each sub-part, homograph entry, sense, or sub-sense we call a level; at any level in an entry, any or all of the constituent parts of dictionary entries may appear. The hierarchical levels of dictionary entries are documented in section [9.2.1 Hierarchical Levels](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DIENHI).

* » [9.2.2 Groups and Constituents](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DIENGP)
* [Home](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html)

#### [¶](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DIENHI)9.2.1 Hierarchical Levels

The outermost structural level of an entry is marked with the elements [entry](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-entry.html) or [entryFree](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-entryFree.html). The [hom](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-hom.html) element marks the subdivision of entries into homographs differing in their part-of-speech. The [sense](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-sense.html) element marks the subdivision of entries and part-of-speech homographs into senses; this element nests recursively in order to provide for a hierarchy of sub-senses of any depth. It is recommended to use the [sense](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-sense.html) element even for an entry that has only one sense to group together all parts of the definition relating to the word sense since this leads to more consistent encoding across entries. All of these levels may each contain any of the constituent parts of an entry. A special case of hierarchical structure is represented by the [re](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-re.html) (related entry) element, which is discussed in section [9.3.6 Related Entries](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPRE). Finally, the element [dictScrap](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-dictScrap.html) may be used at any point in the hierarchy to delimit parts of the dictionary entry which are structurally anomalous, as further discussed in section [9.6 Unstructured Entries](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DIFR).

* [entry](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-entry.html) contains a single structured entry in any kind of lexical resource, such as a dictionary or lexicon.
* [entryFree](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-entryFree.html) (unstructured entry) contains a single unstructured entry in any kind of lexical resource, such as a dictionary or lexicon.
* [hom](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-hom.html) (homograph) groups information relating to one homograph within an entry.
* [sense](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-sense.html) groups together all information relating to one word sense in a dictionary entry, for example definitions, examples, and translation equivalents.

|  |  |
| --- | --- |
| level | gives the nesting depth of this sense. |

* [dictScrap](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-dictScrap.html) (dictionary scrap) encloses a part of a dictionary entry in which other phrase-level dictionary elements are freely combined.

The hierarchic structure of a dictionary entry is enforced by the structures defined in this module. The content model for [entry](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-entry.html) specifies that entries do not nest, that homographs nest within entries, and that senses nest within entries, homographs, or senses, and may be nested to any depth to reflect the embedding of sub-senses. Any of the top-level constituents ([def](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-def.html), [usg](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-usg.html), [form](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-form.html), etc.) can appear at any level (i.e., within entries, homographs, or senses).

* « [9.2.1 Hierarchical Levels](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DIENHI)
* [Home](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html)

#### [¶](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DIENGP)9.2.2 Groups and Constituents

As noted above, dictionary entries, and subordinate levels within dictionary entries, may comprise several constituent parts, each providing a different type of information about the word treated. The top-level constituents of dictionary entries are:

* information about the form of the word treated (orthography, pronunciation, hyphenation, etc.)
* grammatical information (part of speech, grammatical sub-categorization, etc.)
* definitions or translations into another language
* etymology
* examples
* usage information
* cross-references to other entries
* notes
* entries (often of reduced form) for related words, typically called related entries

Any of the hierarchical levels ([entry](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-entry.html), [entryFree](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-entryFree.html), [hom](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-hom.html), and [sense](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-sense.html)) may contain any of these top-level constituents, since information about word form, particular grammatical information, special pronunciation, usage information, etc., may apply to an entire entry, or to only one homograph, or only to a particular sense. The examples below illustrate this point.

The following elements are used to encode these top-level constituents:

* [form](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-form.html) (form information group) groups all the information on the written and spoken forms of one headword.
* [gramGrp](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-gramGrp.html) (grammatical information group) groups morpho-syntactic information about a lexical item, e.g. [pos](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-pos.html), [gen](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-gen.html), [number](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-number.html), [case](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-case.html), or [iType](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-iType.html) (inflectional class).
* [def](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-def.html) (definition) contains definition text in a dictionary entry.
* [cit](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-cit.html) (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](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-usg.html) (usage) contains usage information in a dictionary entry.
* [xr](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-xr.html) (cross-reference phrase) contains a phrase, sentence, or icon referring the reader to some other location in this or another text.
* [etym](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-etym.html) (etymology) encloses the etymological information in a dictionary entry.
* [re](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-re.html) (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](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-note.html) contains a note or annotation.

### 9.3 Top-level Constituents of Entries

This section describes the top-level constituents of dictionary entries, together with the phrase-level constituents peculiar to each.

* the [form](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-form.html) element, which groups orthographic information and pronunciations, is described in section [9.3.1 Information on Written and Spoken Forms](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPFO)
* the [gramGrp](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-gramGrp.html) element, which groups elements for the grammatical characterization of the headword, is described in section [9.3.2 Grammatical Information](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPGR)
* the [def](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-def.html) element, which describes the meaning of the headword, is described in section [9.3.3 Sense Information](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPSE)
* the [etym](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-etym.html) element and its special phrase-level elements are documented in section [9.3.4 Etymological Information](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPET)
* the [cit](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-cit.html) element and its specific applications are described in section [9.3.3 Sense Information](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPSE) and section [9.3.5 Other Information](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPMI)
* the [usg](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-usg.html), [lbl](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-lbl.html), [xr](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-xr.html), and [note](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-note.html) elements are described in section [9.3.5 Other Information](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPMI)
* the [re](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-re.html) element, which marks nested entries for related words, is described in section [9.3.6 Related Entries](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPRE)

#### 9.3.3 Sense Information

Dictionaries may describe the meanings of words in a wide variety of different ways—by means of synonyms, paraphrases, translations into other languages, formal definitions in various highly stylized forms, etc. No attempt is made here to distinguish all the different forms which sense information may take; all of them may be tagged using the [def](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-def.html) element described in section [9.3.3.1 Definitions](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPDE).

As a special case it is frequently desirable to distinguish the provision of translation equivalents in other languages from other forms of sense information; the use of <cit type="translation"> (which groups a translation equivalent with related information such as its grammatical description) for this purpose is described in section [9.3.3.2 Translation Equivalents](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPTR).

* » [9.3.3.2 Translation Equivalents](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPTR)
* [Home](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html)

##### [¶](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPDE)9.3.3.1 Definitions

Dictionary definitions are those pieces of prose in a dictionary entry that describe the meaning of some lexical item. Most often, definitions describe the headword of the entry; in some cases, they describe translated texts, examples, etc.; see <cit type="translation">, section [9.3.3.2 Translation Equivalents](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPTR), and <cit type="example">, section [9.3.5.1 Examples](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DITPEG). The [def](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-def.html) element directly contains the text of the definition; unlike [form](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-form.html) and [gramGrp](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-gramGrp.html), it does not serve solely to group a set of smaller elements. The close analysis of definition text, such as the tagging of hypernyms, typical objects, etc., is not covered by these Guidelines.

##### 9.3.3.2 Translation Equivalents

Multilingual dictionaries contain information about translations of a given word in some source language for one or more target languages. Minimally, the dictionary provides the corresponding translation in the target language; other material, such as morphological information (gender, case), various kinds of usage restrictions, etc., may also be given. If translation equivalents are to be distinguished from other kinds of sense information, they may be encoded using <cit type="translation">. The global xml:lang attribute should be used to specify the target language.

As in monolingual dictionaries, the [sense](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-sense.html) element is used in multilingual dictionaries to group information (forms, grammatical information, usage, translation(s), etc.) about a given sense of a word where necessary. Information about the individual translation equivalents within a sense is grouped using <cit type="translation">. This information may include the translation text (tagged [q](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-q.html) or [quote](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-quote.html)), morphological information ([gen](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-gen.html), [case](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-case.html), etc.), usage notes ([usg](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-usg.html)), translation labels ([lbl](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-lbl.html)), and definitions ([def](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-def.html)).When bibliographic data is provided, the [quote](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-quote.html) element should be used.

* [cit](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-cit.html) (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.
* [lbl](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-lbl.html) (label) contains a label for a form, example, translation, or other piece of information, e.g. abbreviation for, contraction of, literally, approximately, synonyms:, etc.

#### 9.3.4 Etymological Information

The element [etym](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-etym.html) marks a block of etymological information. Etymologies may contain highly structured lists of words in an order indicating their descent from each other, but often also include related words and forms outside the direct line of descent, for comparison. Not infrequently, etymologies include commentary of various sorts, and can grow into short (or long!) essays with prose-like structure. This variation in structure makes it impracticable to define tags which capture the entire intellectual structure of the etymology or record the precise interrelation of all the words mentioned. It is, however, feasible to mark some of the more obvious phrase-level elements frequently found in etymologies, using tags defined in the core module or elsewhere in this chapter. Of particular relevance for the markup of etymologies are:

* [etym](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-etym.html) (etymology) encloses the etymological information in a dictionary entry.
* [lang](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-lang.html) (language name) contains the name of a language mentioned in etymological or other linguistic discussion.
* [date](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-date.html) contains a date in any format.
* [mentioned](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-mentioned.html) marks words or phrases mentioned, not used.
* [gloss](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-gloss.html) identifies a phrase or word used to provide a gloss or definition for some other word or phrase.
* [pron](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-pron.html) (pronunciation) contains the pronunciation(s) of the word.
* [usg](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-usg.html) (usage) contains usage information in a dictionary entry.
* [lbl](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-lbl.html) (label) contains a label for a form, example, translation, or other piece of information, e.g. abbreviation for, contraction of, literally, approximately, synonyms:, etc.

As in other prose, individual word forms mentioned in an etymological description are tagged with [mentioned](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-mentioned.html) elements. Pronunciations, usage labels, and glosses can be tagged using the [pron](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-pron.html), [usg](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-usg.html), and [gloss](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-gloss.html) elements defined elsewhere in these Guidelines. In addition, the [lang](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-lang.html) element may be used to identify a particular language name where it appears, in addition to using the xml:lang attribute of the [mentioned](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-mentioned.html) element.

##### 9.3.5.3 Cross-References to Other Entries

Dictionary entries frequently refer to information in other entries, often using extremely dense notations to convey the headword of the entry to be sought, the particular part of the entry being referred to, and the nature of the information to be sought there (synonyms, antonyms, usage notes, etymology, an illustration, etc.)

Cross-references may be tagged in dictionaries using the [ref](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-ref.html) and [ptr](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-ptr.html) elements defined in the core module (section [3.6 Simple Links and Cross-References](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/CO.html#COXR)). In addition, the [xr](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-xr.html) element may be used to group all the information relating to a cross-reference.

* [xr](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-xr.html) (cross-reference phrase) contains a phrase, sentence, or icon referring the reader to some other location in this or another text.
* [ref](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-ref.html) (reference) defines a reference to another location, possibly modified by additional text or comment.
* [ptr](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-ptr.html) (pointer) defines a pointer to another location.
* [lbl](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-lbl.html) (label) contains a label for a form, example, translation, or other piece of information, e.g. abbreviation for, contraction of, literally, approximately, synonyms:, etc.

As in other types of text, the actual pointing element (e.g. [ref](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-ref.html) or [ptr](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-ptr.html)) is used to tag the cross-reference target proper (in dictionaries, usually the headword, possibly accompanied by a homograph number, a sense number, or other further restriction specifying what portion of the target entry is being referred to). The [xr](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-xr.html) element is used to group the target with any accompanying phrases or symbols used to label the cross-reference; the cross-reference label itself may be tagged as a [lbl](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-lbl.html) or may remain untagged. Both of the following are thus legitimate:

**glee** … Compare **madrigal** (sense 1) CED

<entry>  
 <form>  
  <orth>glee</orth>  
 </form>  
 <xr>Compare <ptr target="#madrigal.1"/>  
 </xr>  
</entry>  
<entry xml:id="madrigal.1">  
 <form>  
<!-- ... -->  
 </form>  
</entry>

**hostellerie** Syn. de hôtellerie (sens 1). DNT

<xr type="syn">  
 <lbl>Syn. de</lbl>  
 <ref>hôtellerie (sens 1)</ref>.  
</xr>

In addition to using, or not using, [lbl](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-lbl.html) to mark the cross-reference label, the two examples differ in another way. The former assumes that the first sense of madrigal has the identifier madrigal.1, and that the specific form of the reference in the source volume can be reconstructed, if needed, from that information. The latter does not require the first sense of ‘hôtellerie’ to have an identifier, and retains the print form of the cross-reference; by omitting the target attribute of the [ref](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-ref.html) element, however, the second example does assume implicitly either that some software could usefully parse the phrase tagged as a [ref](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-ref.html) and find the location referred to, or else that such processing will not be necessary.

The type attribute on the pointing element or on the [xr](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-xr.html) element may be used to indicate what kind of cross-reference is being made, using any convenient typology. Since different dictionaries may label the same kind of cross-reference in different ways, it may be useful to give normalized indications in the type attribute, enabling the encoder to distinguish irregular forms of cross-reference more reliably:

**rose**2 … vb. the past tense of **rise**. CED

<entry n="2">  
 <form>  
  <orth>rose</orth>  
 </form>  
 <xr type="inflectedForm">  
  <lbl>the past tense of</lbl>  
  <ref target="#rise">rise</ref>  
 </xr>  
</entry>  
<!-- ... -->  
<entry xml:id="rise">  
 <form>  
  <orth>rise</orth>  
 </form>  
<!-- main entry for "rise" as verb -->  
</entry>

from cross-references for synonyms and the like:

**antagonist** … syn see **adverse** W7

<xr type="synonym">  
 <lbl>syn see</lbl>  
 <ref target="#adverse">adverse</ref>  
</xr>  
<!-- ... -->  
<entry xml:id="adverse">  
 <form>  
  <orth>adverse</orth>  
 </form>  
<!-- list of synonyms for "adverse" -->  
</entry>

Strictly speaking, the reference above is not to the entry for adverse, but to the list of synonyms found within that entry. In some cases, the cross-reference is to a particular subset of the meanings of the entry in question:

**globe** …V. **armillaire** (sphère) PR

<xr>V. <ref target="#armillaire">armillaire</ref>  
 <lbl type="sense-restriction">sphère</lbl>  
</xr>

Cross-references occasionally occur in definition texts, example texts, etc., or may be free-standing within an entry. These may typically be encoded using [ref](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-ref.html) or [ptr](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-ptr.html), without an enclosing [xr](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-xr.html). For example:

**entacher** … Acte entaché de nullité, contenant un vice de forme ou passé par un incapable\*. DNT

The asterisk signals a reference to the entry for incapable.

<def>contenant un vice de forme ou passé par un <ptr target="#incapable"/>.</def>

In some cases, the form in the definition is inflected, and thus [ref](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-ref.html) must be used to indicate more exactly the intended target, as here:

**justifier** …4. IMPRIM Donner a (une ligne) une longeur convenable au moyen de blancs (2, sens 1, 3). DNT

<sense n="4">  
 <usg type="dom">imprim</usg>  
 <def>Donner a (une ligne) une longeur convenable au moyen de  
 <ref target="#blanc-2.1.3">blancs (2, sens 1, 3)</ref>  
 </def>  
</sense>  
<entry xml:id="blanc" n="2">  
<!-- ... -->  
 <sense n="1">  
<!-- ... -->  
  <def xml:id="blanc-2.1.3">...</def>  
<!-- ... -->  
 </sense>  
<!-- ... -->  
</entry>