GUIDELINES FOR THE MORPHOLOGICAL AND SYNTACTIC ANNOTATION OF OLD ENGLISH

Kristin Bech
# TABLE OF CONTENTS

1. **Introduction**  

2. **Morphology**  
   2.1 Word classes  
      2.1.1 Verb  
      2.1.2 Noun, common  
      2.1.3 Noun, proper  
      2.1.4 Pronoun, personal  
      2.1.5 Pronoun, interrogative  
      2.1.6 Pronoun, indefinite  
      2.1.7 Pronoun, demonstrative  
      2.1.8 Pronoun, possessive  
      2.1.9 Quantifier  
      2.1.10 Adjective  
      2.1.11 Adverb, general  
      2.1.12 Adverb, interrogative  
      2.1.13 Preposition  
      2.1.14 Conjunction  
      2.1.15 Subjunction  
      2.1.16 Infinitive marker  
      2.1.17 Interjection  
      2.1.18 Foreign word  
      2.1.19 Other possible word classes  
   2.2 Lemmatization  
      2.2.1 *beon/wesan*  
      2.2.2 The prefix *ge-*  
      2.2.3 Syncretism  
      2.2.4 Place names and proper names  
      2.2.5 Participles  
      2.2.6 Comments on various problem lemmas  
   2.3 Latin case endings
### Syntax

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Sentence division</td>
<td>20</td>
</tr>
<tr>
<td>3.2</td>
<td>Verb phrases with ‘modal auxiliary’ + infinitive</td>
<td>21</td>
</tr>
<tr>
<td>3.3</td>
<td>Accusative with infinitive</td>
<td>22</td>
</tr>
<tr>
<td>3.4</td>
<td><em>hatan</em> + infinitive</td>
<td>24</td>
</tr>
<tr>
<td>3.5</td>
<td>Verb phrases with <em>be</em> + past participle</td>
<td>28</td>
</tr>
<tr>
<td>3.6</td>
<td>Verb phrases with <em>be</em> + present participle</td>
<td>30</td>
</tr>
<tr>
<td>3.7</td>
<td>Adverbs and preposition phrases with the verb <em>be</em></td>
<td>36</td>
</tr>
<tr>
<td>3.8</td>
<td>Verb phrases with <em>habban</em> + participle</td>
<td>37</td>
</tr>
<tr>
<td>3.9</td>
<td>Subject complements</td>
<td>38</td>
</tr>
<tr>
<td>3.10</td>
<td>Relative clauses</td>
<td>39</td>
</tr>
<tr>
<td>3.10.1</td>
<td><em>pe</em> relatives</td>
<td>39</td>
</tr>
<tr>
<td>3.10.2</td>
<td><em>se</em> relatives</td>
<td>40</td>
</tr>
<tr>
<td>3.10.3</td>
<td><em>se pe</em> relatives</td>
<td>42</td>
</tr>
<tr>
<td>3.10.4</td>
<td><em>eall pe</em> constructions</td>
<td>45</td>
</tr>
<tr>
<td>3.10.5</td>
<td>Headless (nominal) relatives</td>
<td>46</td>
</tr>
<tr>
<td>3.11</td>
<td>Comparison</td>
<td>50</td>
</tr>
<tr>
<td>3.12</td>
<td>Appositions</td>
<td>55</td>
</tr>
<tr>
<td>3.13</td>
<td>Complex subjunctions</td>
<td>56</td>
</tr>
<tr>
<td>3.14</td>
<td>Complex conjunction</td>
<td>57</td>
</tr>
<tr>
<td>3.15</td>
<td><em>both ... and</em></td>
<td>58</td>
</tr>
<tr>
<td>3.16</td>
<td>Negation</td>
<td>58</td>
</tr>
<tr>
<td>3.17</td>
<td>Impersonal constructions</td>
<td>58</td>
</tr>
<tr>
<td>3.18</td>
<td>Existential sentences and existential <em>there</em></td>
<td>59</td>
</tr>
<tr>
<td>3.19</td>
<td>Expletive <em>hit</em></td>
<td>60</td>
</tr>
<tr>
<td>3.20</td>
<td>Personal pronoun as reflexive</td>
<td>60</td>
</tr>
<tr>
<td>3.21</td>
<td><em>self</em></td>
<td>61</td>
</tr>
<tr>
<td>3.22</td>
<td><em>that is</em> (<em>i.e.</em>).</td>
<td>61</td>
</tr>
<tr>
<td>3.23</td>
<td>Partitive genitive</td>
<td>62</td>
</tr>
<tr>
<td>3.24</td>
<td>Direct and indirect questions</td>
<td>62</td>
</tr>
<tr>
<td>3.25</td>
<td><em>swa hwæt swa, swa hwilc swa</em></td>
<td>62</td>
</tr>
<tr>
<td>3.26</td>
<td>ARG</td>
<td>62</td>
</tr>
<tr>
<td>3.27</td>
<td>NARG</td>
<td>63</td>
</tr>
<tr>
<td>3.28</td>
<td>Two OBL arguments</td>
<td>63</td>
</tr>
<tr>
<td>3.29</td>
<td>Infinitive of intention/purpose = XADV</td>
<td>64</td>
</tr>
<tr>
<td>3.30</td>
<td>Focus adverb</td>
<td>65</td>
</tr>
<tr>
<td>3.31</td>
<td>Absolute constructions</td>
<td>65</td>
</tr>
<tr>
<td>3.32</td>
<td>Source and goal arguments</td>
<td>68</td>
</tr>
</tbody>
</table>
3.33  butan  68
3.34  PARPRED  69
3.35  Topicalization  70
3.36  FORDAM as conjunction  70
3.37  SOLICE and WITODLICE  71

4  References  73
1 INTRODUCTION
These guidelines are not meant to be complete guidelines to the annotation of the Old English texts. Knowledge of the basic syntactic (dependency grammar) annotation is taken for granted. As regards some of the basic issues, as well as some of the more complex issues, we refer to the PROIEL guidelines for old Indo-European languages, which ISWOC generally follows, though with language-specific exceptions. Reference is also made to the YCOE guidelines for Old English and to the Menotec guidelines for Old Norse. Our guidelines, in addition to providing an overview of the annotation method for some important issues, discuss matters and problems that arose as the annotation progressed, and present the chosen solutions. The numbers in parentheses refer to the sentence ID of the relevant examples. The examples can thus easily be looked up in the corpus if they are not given in full here.

PROIEL: Pragmatic Resources in Old Indo-European Langages

PROIEL guidelines: http://folk.uio.no/daghaug/syntactic_guidelines.pdf

YCOE: York-Toronto-Helsinki Parsed Corpus of Old English Prose

YCOE guidelines: http://www-users.york.ac.uk/~lang22/YCOE/YcoeHome.htm

Menotec: Medieval Norwegian Text Corpus

Menotec guidelines: https://bells.uib.no/bells/issue/view/157
2 MORPHOLOGY
2.1 WORD CLASSES
As regards word classes, the problems we experienced mostly had to do with the ‘small’ function words, such as prepositions, conjunctions, pronouns, determiners and quantifiers. Below, the word classes are presented, and comments on various problem words are made, where relevant.

2.1.1 Verb
Verbs are generally unproblematic morphologically, except participles (see also 3.5 and 3.6.), and except the indicative/subjunctive distinction.

It is not always easy to distinguish between verbal and adjectival uses of the OE participles, because it is not necessarily the case that they are adjectival when they are inflected, and verbal when they are non-inflected. We annotate in the same way as the YCOE corpus (see ‘Case on participles’ in their guidelines). In other words, participles which are part of the main verb sequence, and participles in small clauses are not annotated for case, number, gender and definiteness unless they have overt case. Participles which are modifying or attributive are labeled for case even if it is not overt, except ‘naming participles’, e.g. gehaten, which are only labelled for case if it is overt.

Compare 118755, Swa hwilc man swa me Apollonium lifigendne to gebringð, ic him gife fifti punda goldes ‘Whichever man brings Apollonius living [alive] to me, I will give him fifty pounds of gold’, in which the participle and the proper name are marked for accusative, to 118783 Swa hwilc man swa ðe lifigende to him bringð, onfo se fiftig punda goldes ‘Whichever man brings you living [alive] to him, he will receive fifty pounds of gold’, in which case is not clearly marked, although ðe and lifigende are most probably accusatives.

As regards indicative and subjunctive, there is syncretism in some forms. The annotation practice is that we leave the mood ‘unspecified’ if the indicative and the subjunctive are identical, even if we know that a form is in all likelihood either indicative or subjunctive. It is a practice that can be questioned, but we follow the practice of the YCOE corpus in this matter.

2.1.2 Noun, common
No particular problems.

2.1.3 Noun, proper
No particular problems.

2.1.4 Pronoun, personal
The third person plural pronoun hi is masculine, feminine, or neuter (common). If the gender is unambiguous in the context, it is annotated as either m or f or n. If it is not unambiguous (a crowd, for example), it is annotated as m / f / n.
2.1.5 Pronoun, interrogative

Hwæt: There are three lemmas for hwæt: the interrogative pronoun, the interjection and the indefinite pronoun (see 2.1.6). An example of hwæt as interrogative pronoun is shown in (1).

(1) Nat ic hwæt he besorgað (118920)
not-know I what he bemourns
‘I do not know what he is sorrowful for’

Hwy (hwi) is originally the instrumental case of hwæt. However, I have chosen to label hwy as a non-inflected interrogative adverb in the morphological annotation, since hwy is usually an ADV in the syntactic annotation, and since they are tagged differently in the information structure annotation: the interrogative and indefinite pronoun hwæt usually gets a NONSPEC tag, whereas hwy is not tagged.

Hwilc can be an interrogative pronoun or an indefinite pronoun (see 2.1.6). An example of the interrogative use is found in (2).

(2) Hwilc eower is forliden? (119050)
Who you-GEN is shipwrecked?
‘Who of you is shipwrecked?’

Notice that we do not distinguish between interrogative pronouns and interrogative determiners. Thus in For hwilcum intingum? ‘For what reasons?’ (118803), hwilcum is classified as an interrogative pronoun.

Hwæðer can be an interrogative pronoun or a conjunction, as in (3):

(3) Hwæðer is eðre to secgenne to þam laman, þe synd ðine synna forgysfene, hwæðer þe cweðan, aris, nim ðin bed & ga? (102370)
whether is easier to say to say to the lame, you-DAT are your sins forgiven, whether þe say, arise, take your bed and go?
‘Whether it is easier to say to the sick of the palsy, Thy sins are forgiven, or to say, arise, take your bed and walk?’

The first hwæðer is an interrogative pronoun meaning ‘what’, and the second hwæðer, in combination with þe, has been analyzed as a conjunction, meaning ‘or’. Hwæðer can also be a subjunction, as in (4):

(4) He sæde þet he æt sumum cirre wolde fandian hu longe þat land norþryhte læge, ofþe hwæðer ænig mon be norðan þæm westenne bude (115571)
he said that he at some time wanted find how far that land northwards lay, or whether any man to north that wilderness lived
‘He said that he at some occasion wanted to find out how far north the land extended, or whether any man north of the wilderness lived’
The YCOE guidelines call all these *wh*-words, and they distinguish between *wh*-pronouns, *wh*-adjectives, *wh*-adverbs, *hwæðer* and *swæðer/swa hwæðer*.

### 2.1.6 Pronoun, indefinite

*Hwæt* also occurs in non-interrogative contexts, as in (5), and is in such cases classified as an indefinite pronoun.

(5)  

*Gif him swa hwæt swa ðu wille* (118974)  

give him so what so you will  

‘Give him whatever you wish’

*Hwilc* is analyzed as an indefinite pronoun in constructions such as (6):

(6)  

*sва hwylc hus swa ge ingað, wuniad þar od þæt ge utgan* (102654)  

so which house so you ingo, remain there until that you outgo  

‘Whichever place you enter, remain there until you depart (from that place)’

Note that we do not distinguish between indefinite pronouns and indefinite determiners, so in (6), even though *hwylc* is not the head, but an ATR to *hus* in the syntax, it gets the label ‘indefinite pronoun’.

Also, *man* ‘one’ occurs once in a while, and is labelled an indefinite pronoun.

### 2.1.7 Pronoun, demonstrative

We do not distinguish between demonstrative determiners and demonstrative pronouns; i.e., we do not have ‘demonstrative determiner’ word class, only ‘demonstrative pronoun’. This is because the difference is not morphological, but syntactic, and it becomes apparent on the syntactic level. Thus in *se mann* ‘the man’, where *se* is a determiner, *se* is an ATR to the head *mann*, whereas if *se* is on its own as a pronoun, it constitutes the head.

It happens that the demonstrative determiner *pe* is used instead of *se*, cf. e.g. 118981. In these cases, *pe* is analyzed as ‘non-inflecting’ morphologically, but the lemma form is *se*. In 118981, *pe* may be a foreshadowing of the change to come, or it may be a scribal error, or it may be a northern dialect form, or it may be an error made in the edition of the manuscript, either in the book or in the electronic version. Elsewhere in *Apollonius*, *se* is used.

In 118677, there is an instance of *pe* used before a comparative: *pe lenge* ‘the longer’. Most instances of *pe* occur in relative clauses, where they are analyzed as a subjunction. However, in this case as well, it was classified as a non-inflecting demonstrative pronoun.

### 2.1.8 Pronoun, possessive

The possessive pronouns *his, hire* (sg.f.) and *hira* (pl.) have been tagged as non-inflecting possessive pronouns, since they never get case endings. The other possessive pronouns have
been tagged as inflecting, since they get case. Note also that the possessive pronouns are separate lemmas. For example, min has the lemma min, and is not regarded as the genitive form of the personal pronoun ic. The lemma forms for the inflecting possessive pronouns are thus: min, þin, ure, (uncer), eower. Note that the possessive pronouns are also tagged for person and number. These are superfluous categories for this word class, and should probably have been removed from the system, but we thought of it too late. In the Menotec corpus, the third person forms corresponding to OE his, hire, hira are regarded as personal pronouns; thus the analysis for Old Norwegian differs from our analysis in this matter.

2.1.9 Quantifier
This is a word class which, in retrospect, should probably have been handled differently. The corpus texts were imported from the YCOE corpus, by means of a script that also imported their morphological annotation. As regards quantifiers, the classification might have been more fine-grained, taking syntactic function into consideration to a greater extent (cf. also Ann Taylor, personal communication, who says that she is not particularly satisfied with the way quantifiers were handled in the YCOE corpus).

For example, micel and lytel are classified as quantifiers throughout, even when they are adjectival, as in YCOE, cf. their comment:

Note especially that (UN)LYTEL and MICEL are consistently tagged Q, even when their meaning is more adjectival than quantificational, i.e. when LYTEL is better interpreted as small and MICEL as great. Distinguishing between the two readings can be quite difficult, especially with plural nouns, and we have not attempted to do so, despite the fact that it creates infelicitous readings in some cases.

(7) is an example in which mycele is adjectival and modified by the adverb swa.

(7) & þonne hit asawen bið hit astihþ, & bið ealra wyrta maest & heafð swa mycele bogas þæt heofenes fugelas eardian magon under his sceade. (102531)
and when it sown is it grows and is all-GEN herbs-GEN greatest and has so great branches that air’s birds lodge may under its shadow
‘and when it is sown, it grows up, and becomes greater than all herbs, and has such great branches that all the birds of the air may lodge in its shadow’

---

But see comments on URE and EOWER in the YCOE guidelines. They sometimes fail to agree with the following noun, and are in such cases not tagged for case, but as ‘non-inflecting’ in our system.
Further comments on quantifiers:

118781, 118822: *Micel* is used adverbially, but it is nevertheless a quantifier in the morphology (see also YCOE guidelines). Cf. also 107871. Notice the placement of *mycelum*, as a dependent on the participle (normally, adverbials are dependent on the main verb). Cf. also *eal* in 107844, 119063, 119068.

118732 (e.g.): *Naht* (lemma *nawiht*) is not easy to categorize, but we have followed YCOE and tagged it as a quantifier, even in the cases where it is used adverbially or nominally, or as a modifier.

118968 (e.g.): *fela* ‘many’ is rarely inflected, and is tagged as non-inflecting.

118835, 118836: The function of *eal* may sometimes be difficult to determine – whether it is an adverbial meaning ‘entirely’, or an attributive element. In these cases, they have been taken to modify the noun.

118741

118680: *Mare* is used adverbially but tagged as a quantifier.

Cardinal numerals have been tagged as quantifiers.
In (8), the comparative form of *micel, ma*, is used in the expression *ma þonne*, which approaches the meaning ‘rather than’ and is what Quirk et al. (1985: 761, 982) term a ‘quasi-coordinator’, i.e. an item that sometimes behaves like a coordinator and sometimes like a subjunction or a preposition. In this case, we have taken it to be a coordinator – ‘he took a merchant’s name rather than a giver’s’, and analyzed *ponne* (here taken as a preposition in the morphology) as the coordinating element and *ma* as an ADV to *ponne*.

(8)  *Hwæt ða Apollonius forlet his þone wurðfullan cynedom and mangeres naman par genam ma þonne gifendes* (118822)

### 2.1.10 Adjective

*Ana* ‘only’, ‘alone’ has been analyzed as an adjective throughout, although there is a difference between *Only he went into the room*, and *He went alone into the room*. In the former case, *ana* is analyzed as an ATR to *he*; in the latter case, *ana* is analyzed as an XADV, i.e. an adverbial element with an external subject. The two analyses are shown in (9) and (10), respectively.

(9)  *and Apollonius ana swigode.* (1118952)

and Apollonius only was-silent

‘and only Apollonius was silent’

(10) *Lareow, hwi gæst ðu ana?* (119029)

Master, why go you alone?

‘Master, why do you go alone?’

Concerning quantifiers used as adjectives, see section 2.1.9 above.

Ordinal numerals are tagged as adjectives.

### 2.1.11 Adverb, general

*Swilce* can be both an adjective and an adverb, and as adverb it usually occurs in the construction *eac swilce*.

*Hwilum* is probably the dative plural of *hwil*, used adverbially. We have, however taken it as an adverb in the morphology (115626). YCOE is not consistent.
Ham is both an adverb and a common noun. In expressions such as *ham gewendan* ‘go home’, *ham* is taken to be adverb, while in *fram his agnum ham* ‘from his own home’, it is taken to be a noun.

2.1.12 Adverb, interrogative
In this word class are found the interrogative adverbs *hu, hwy, hwar*.

2.1.13 Preposition
No particular problems.

2.1.14 Conjunction
No particular problems. The main conjunctions are *and* and *ac*. See the comment on *hwæder* in 2.1.5.

2.1.15 Subjunction
Here it should be noted that the relative particle *pe* in Old English, like the Old Norse relative particles *er* and *sem*, is taken to be a subjunction and not a pronoun, since it does not have any case, number or person agreement. However, unlike Old Norse, Old English also has relative clauses that are introduced by a demonstrative determiner (*se*) functioning as relative pronoun. This means that relative clauses introduced by *pe* and relative clauses introduced by a form of *se* are analyzed differently in the syntactic analysis, cf section 3.10. There is also the complicated matter of complex subjunctions, and I refer to section 3.13 below for a discussion of these.

2.1.16 Infinitive marker
No problem here – the infinitive marker is *to*.

2.1.17 Interjection
Common interjections are *eala* and *hwæt*. Walkden (2013) discusses *hwæt*, and concludes that it is in fact not an interjection, and thus not extra-clausal, but an underspecified wh-pronoun introducing an exclamative clause. This may well be the case, but for the time being, we keep to the traditional analysis of *hwæt* as an interjection.

2.1.18 Foreign word
Since the old texts are often translations (from Latin, mostly), it happens that some foreign words are kept in the text, and these are tagged as ‘foreign word’. See 2.2.6 for a comment on the use of the Latin conjunction *et* ‘and’.

2.1.19 Other possible word classes
YCOE does not tag relative adverbs specifically as such (they tag them as locative adverbs); neither do we. In our application, we do not distinguish between semantic types of adverbs (locative, temporal, etc), hence relative adverbs are simply classified as ‘adverbs’. An example is (11):

---

*Ham* is both an adverb and a common noun. In expressions such as *ham gewendan* ‘go home’, *ham* is taken to be adverb, while in *fram his agnum ham* ‘from his own home’, it is taken to be a noun.

2.1.12 Adverb, interrogative
In this word class are found the interrogative adverbs *hu, hwy, hwar*.

2.1.13 Preposition
No particular problems.

2.1.14 Conjunction
No particular problems. The main conjunctions are *and* and *ac*. See the comment on *hwæder* in 2.1.5.

2.1.15 Subjunction
Here it should be noted that the relative particle *pe* in Old English, like the Old Norse relative particles *er* and *sem*, is taken to be a subjunction and not a pronoun, since it does not have any case, number or person agreement. However, unlike Old Norse, Old English also has relative clauses that are introduced by a demonstrative determiner (*se*) functioning as relative pronoun. This means that relative clauses introduced by *pe* and relative clauses introduced by a form of *se* are analyzed differently in the syntactic analysis, cf section 3.10. There is also the complicated matter of complex subjunctions, and I refer to section 3.13 below for a discussion of these.

2.1.16 Infinitive marker
No problem here – the infinitive marker is *to*.

2.1.17 Interjection
Common interjections are *eala* and *hwæt*. Walkden (2013) discusses *hwæt*, and concludes that it is in fact not an interjection, and thus not extra-clausal, but an underspecified wh-pronoun introducing an exclamative clause. This may well be the case, but for the time being, we keep to the traditional analysis of *hwæt* as an interjection.

2.1.18 Foreign word
Since the old texts are often translations (from Latin, mostly), it happens that some foreign words are kept in the text, and these are tagged as ‘foreign word’. See 2.2.6 for a comment on the use of the Latin conjunction *et* ‘and’.

2.1.19 Other possible word classes
YCOE does not tag relative adverbs specifically as such (they tag them as locative adverbs); neither do we. In our application, we do not distinguish between semantic types of adverbs (locative, temporal, etc), hence relative adverbs are simply classified as ‘adverbs’. An example is (11):
and one day on early-morning when he of sleep woke, he broke into the chamber where she in lay...

‘and one day, in the morning, when he woke from sleep, he broke into the chamber wherein she lay’.

It would have been possible to tag these as relative adverbs, since that word class exists in the corpus application, and with hindsight, it might have been better to do so. However, as it is, þær (lemma þæ) is analyzed as a general adverb, so we do not make a distiction between the meanings ‘there’ and ‘where’ for þær. It is possible to find these instances by searching for the lemma þær in an APOS or ATR structure (non-restrictive or restrictive relative clause).

2.2 LEMMATIZATION

The lemma form is the first (main) entry in the digital edition of Bosworth & Toller’s (B&T) Anglo-Saxon Dictionary: [http://bosworth.ff.cuni.cz/](http://bosworth.ff.cuni.cz/). Clark Hall’s (CH) A Concise Anglo-Saxon Dictionary has in some cases been used in addition, e.g. to find out whether a participle form used adjectivally has a related verb listing or not. It may have one in one of the dictionaries, but not in the other. It also happens, in very rare cases, that a word is listed in Clark Hall, but not in B&T. An example is the adverb that occurs in 33802, nāþelæs, which does not occur in B&T (in any spelling variety). It was therefore lemmatized according to CH: nāþylæs. For some reason, B&T has variation in the use of eth (ð) and thorn (þ) in the lemmas. For example, swiðe is lemmatized with eth, and wiþ with thorn. This variation has been kept in our corpus, but it is possible that we should just have lemmatized them all with the more common thorn, for simplicity’s sake.

2.2.1 beon/wesan

For the verb beon/wesan, the lemma form is wesan or beon, depending on the form of the verb. For the forms eom, eart, is, sindon, sy, syn, the lemma form is wesan.

2.2.2 The prefix ge-

Old English verbs may have the prefix ge- in all forms. Thus, both winnan, and gewinnan may mean ‘make war, fight’, but gewinnan has an additional perfective sense ‘get by fighting’, and therefore the prefix is often used in past participles (Mitchell & Robinson, 1995: 58). In other words, the past participle of winnan can be either wunnen or gewunnen. The problem is what to do with the lemma in sentences with participles like gewinnan. Is the lemma winnan or gewinnan? (B&T lists winnan and gewinnan as two different entries.) We did as follows: For forms other than the participle, the lemma is given according to whether ge- is present or not. For example, if the form is geworhte (past tense), the lemma is gewyrce, and if the form is worht, the lemma is wyrce. For participle forms with ge- the lemma always includes ge- as well, since we do not know if it is a participle of a prefixed or a non-prefixed verb. (The opposite solution could also have been chosen, i.e. to have the non-prefixed infinitive as lemma in all cases, but since ge- adds to the meaning of the word, it was felt that the lemma should also include it.)
2.2.3 Syncretism

Two different lemmas may be identical in form. Usually, they belong to different word classes, so distinguishing between them and finding them in queries is no problem. However, it occurs that there are two identical lemmas in the same word class. An example is the common noun *scip*, which usually means ‘ship’, but there is also a lemma *scip* that means ‘patch’ (102400). We have not distinguished between them in the morphology, but it is possible to do so by marking them *scip*\#1 and *scip*\#2, for example. Since they are not distinguished, any search looking for *scip* ‘ship’ will have to go through the examples and filter out the ‘patch’ meanings manually.

It happens quite often that there is syncretism between preterite and subjunctive forms of the verb. In such cases, the mood is left as ‘unspecified’ in the morphology, even in those cases where we are quite sure from the sentence type or context which form it is.

There is also case syncretism for nouns and pronouns; for example, accusative and dative forms may be identical. If it is not possible to determine which case it is, e.g. after a preposition, it is left unspecified. If the dictionary lists other examples where one or the other case form unambiguously occurs, case is determined on the basis of that. An argument for which it is difficult to specify case gets the syntactic function ARG, even when it may be quite clear from the context whether it is OBJ or OBL.

2.2.4 Place names and proper names

Place names and proper names are often difficult to lemmatize because they do not always occur in their nominative form in the texts, and because Latin case endings are sometimes used, which makes it even more difficult to envisage what the nominative form might be (cf. 2.3). In addition, place names are often not listed in the dictionaries. Below are listed some place names and proper names as they are lemmatized in this corpus.

**Names of peoples, tribes:** *Athene, Beormas, Cantware, Cornwealas, Crece* (though *Crecas* actually occurs more often in the texts), *Defenas, Deniscan, Eald-Seaxe, East-Engle, East-Francon, East-Seaxe, Egipte, Francon, Frysa, Judeas, Lidwiccas, Langbeardas, Macedonie, Mirce, Norpan-hymbre* (*Norp-hymbre* also occurs as an entry in B&T, but we have used *Norpan-hymbre* as the lemma throughout), *Norp-Walas, Norp-Wealh-cynn, Peohtas, Perse, Seaxe, Scottas, Sumersæte, Suhrige, Suþ-Seaxe, Thebane, Walas, Wealh, West-Wealas, West-Seaxe* (the form that often occurs is *Wesseaxna*, a genitive plural), *Wilsætan*. See also Quirk & Wrenn, paragraph 30.

As regards England at this time, the names of peoples are usually also used as the place names, cf. e.g. *Angelcynn*. In B&T, these are sometimes listed with gender information – e.g. *East-Seaxe* is masculine plural, and sometimes without – e.g. *Mirce* is just listed as a plural. The annotation follows the information given in B&T. Sometimes it is possible to establish case even if the gender is unknown, e.g. *Miercna cyning* ‘king of the Mercians’. We know in this case that *Miercna* must be a genitive plural, or *Persum*, which must be a dative plural.
Bosworth & Toller is not always consistent in providing the gender information. For example, *Norþ-Walas* is listed as a plural, with no gender specification, whereas *West-Wealas* (with that spelling) is listed as a masculine plural. And *Walas* is not listed at all. It is obvious that *Norþ-Walas* must be a masculine, since it also occurs in the accusative as *Norþ-Walas*. In extremely obvious cases like this, we have annotated the token with gender.

**Place names:** *Affrica, Alexandria, Antiochia, Asia, Capharnaum, Cartagine* (the nominative form is unknown), *Cesarea, Cilicia, Cîönus, Creca, Dalmanûda, Decapolis, Egipte* (Ælfric form: *Ægyfto (lande)*, but lemmatized *Egipte*). See also names of peoples above), *Friga, Galilea, Hieraสนามorum* (this is a case ending, but it is difficult to know what the nominative form might be, as it only occurs in this form in the text), *Macedonia, Oliuete* (Mount of Olives: *Oliuetes Dune; oliuete* is not in any dictionary, so the lemma had to be created. In the other languages (Latin, Greek, French) ‘olives’ is taken to be a common noun), *Perse* (see names of peoples above), *Siria, Tharsus, Thracia, Tirus, Rom, Romeburh, Rodôs*.

Anglo-Saxon Chronicles: *Aler, Andred, Apuldre, Æscesdun, Æðeling, Basengum, Beamfleot, Bune* (Bologne), *Cantwarburg, Cariei* (Chezy, France. No dictionary entry for this, so the form was lemmatized as it was), *Carrum, Cippanham, Cirenceaster, Cisseceaster, Clofeshoh, Coln, Cundöph (Conde)*, *Cynemæresford, Cwatbrycg, Defena scir, Dor, Embene, Englafeld, Eoforvicceaster, Eçandun, Exanceaster* (also spelled Escanceaster), *Fearnham, Frandcland, Fullanham, Gend, Hamtunscir, Hibernia, Hreopedun, Hrofasceaster, Hambre (river), Iglea, Ione* (river Yonne, France. No dictionary entry. The form in the text is *Ionan*. I set the lemma to *Ione*, as it is a river, and rivers are usually feminine), *Legaceaster, Limene* (river Limne), *Lindesse, Lunden, Lundenburh, Mæs* (the French river Meuse), *Mæterne* (river Marne. No dictionary entry for this, so I just took the form as it was.), *Mere, Meretun, Middeltun, Pafía, Paris, Pedriding, Rin* (river Rhine), *Sandwic, Sant Lauda* (Saint Lo), *Scald, Scoeburh, Scireburne, Sealhwudu, Sigen* (river Seine), *Sootingaham, Stræcled, Stufemûpa, Sunna* (river Somne), *Suþrige, Swancwich, Temes, Tenet, Æoford, Wendelæs, Werham, Weþmor, Wicganbeorg, Wiht, Wiltun, Wiltunscir, Winburne, Wintanceaster, Wirhealh*.

Place names have been given the tag ‘unspecified’ for gender, unless they consist of a noun that we know has a gender, e.g. *Manceaster* (ceaster is a feminine noun), *Romeburh* (burh is a feminine noun), or unless the place name is specifically listed with gender in B&T. For example, *Rom* is listed in B&T as a feminine noun. Place names in Old English texts often retained the gender of the original text (mostly Latin), and this is also evident from the fact that some place names get Latin case, e.g. *Tharsus – Tharsum* (Latin accusative), *Tirus – Tirum* (Latin accusative) – *Tiro* (Latin dative). In other words, it is in fact often possible to assign gender and case to place names on the basis of Latin, but we decided to nevertheless leave them unspecified, for reasons of consistency.

**Proper names:** *Alpheus, Leuin, Parmenio, Demostanas, Philippus, Iobes (Jupiter), Eugenia, Sergius* (often spelled *Særgius*), *Diana*. In addition, there are the biblical names, which are quite straightforward: *Crist, Iohannes, Simon, Andreas, Bartholomeus, Iacob, Isaias*, etc. The
Anglo-Saxon Chronicle also features many names: Æbbe, Ælfred, Ælle, Æþelbald, Æþelbryht, Æþelberh, Æþelheard, Æþelhelm, Æþelmund, Æþelnoþ, Æþelstan, Æþelswiþ, Æþelwald, Æþelwulf, Æþelwulfing (as a patronym – son of Æþelwulf), Æþered, Æþelred, Anwynd, Bagsecg (also spelled Bachsecg), Baldred, Beocca, Beorhtric, Beorhtulf, Beorhtwulf, Beorngar, Beornhelm, Beormund, Beornwulf, Burghelm, Burgred, Carl, Ceawlin, Cenwulf, Ceolburg, Ceolmund, Ceolnoþ, Ceolwulf, Ceorl, Cuþred, Eadbryht, Eadmund, Eadulf, Eadwine, Ealchere (also spelled Ealhere), Ealhheard, Ealhstan (also spelled Ealchstan), Eanulf, Earnulf, Ecgbryht, Ecgbrefting (as a patronym – son of Ecgbryht), Ecgulf, Feologid, Forþred, Fræna, Hæsten, Halfdene (also spelled Healdene), Heahstan, Godrum, Hareld, Heabryht, Heahmund, Herebryht, Hereferþ, Hloþwig, Huda, Inwær, Iuþytt, Leo, Lucumon, Ludecan, Maccbethu, Maelinmun, Marinus, Muca, Oda, Offa, Ordheh, Osbearn, Oscytel, Osbryht, Osmod, Osric, Oswald, Oswio, Paschalis, Pippen, Plegemund, Rædwald, Roþulf, Sidroc, Stephanus, Swifneh, Swiþulf, Wæoxtan, Wigbryht, Wiglaf, Wigþen, Wiþa, Worr, Wulfheard, Wulfred, Wulfric.

Proper names are masculine or feminine according to the gender of the person named. The name Eugenia in Ælfriec’s Lives of Saints takes the -an declension (i.e. Eugenian in accusative, genitive and dative) so it is not always easy to know what case it has, and consequently whether the syntactic analysis should be OBJ or OBL. If nothing else in the NP can inform us about the case, or the verb in the sentence in which Eugenian occurs is not listed (in Quirk & Wrenn) as a verb that takes a particular case, or the verb is not listed in B&T as taking a particular case, the case is marked as unspecified, and the syntactic function as ARG (see section 3.26). For example, for Heo bæd þa Eugenian þæt... ‘She then bade Eugenia that... (107867), B&T informs us that biddan is followed by an accusative of the person, and therefore Eugenian gets an accusative tag, and is analyzed as an OBJ. In Mæg gehyran se þe wyle be þam halgan mædene Eugenian... ‘He who will may hear about the holy maiden Eugenia’ (107806), Eugenian is an apposition to mædene, which is a dative, and it thus gets a dative tag. In and ealle mid aðe Eugenian forlu... ‘and all belied Eugenia under oath’ (107897), we do not know for certain what the case of Eugenian is (although it is most probably an accusative), therefore the case is marked as ‘unspecified’, and the syntactic function of Eugenian is ARG.

2.2.5 Participles
When participles are used attributively, they strongly resemble adjectives. However, we decided to tag participles as verbs, with the infinitive form as the lemma. In some cases, however, there was no entry for that particular verb, neither in B&T nor in CH. In such cases, the participle was classified as an adjective. An example is sum æþelboren þægn ‘a nobly-born thane’ in 107807, for which a dictionary entry æþelberan does not exist (the verb beran exists, obviously), only æþelboren. It was therefore tagged as an adjective, with the lemma æþelboren. Other examples are cyneboren in 107946 and foresprecena in 90984.
2.2.6 Comments on various problem lemmas

118657: *pa ongeanwinnendan feaman* ‘the struggling damsel’. Usually, when participles are used attributively, the lemma is the verb form. However, in some cases, B&T does not list a verb form – thus there is no lexical entry *ongeanwinnan*. They do, however, list *ongeanwinnende*, so that was the lemma used in this case.

102280: *Lendenu* ‘loins’ is a plural noun (there is no sg. lemma), and the lemma is therefore this form.

116658: For Roman numerals, the lemma is spelled out in letters. Notice that the lemma is given as *twentigsix*, and not as *six-and-twentig*, which is probably the manner in which they would have said it at that time. See also e.g. 91069. Notice that the numeral CCXX (120) is *hundtwelftig* not *hundtwentig*, cf. 91072. The numeral 100 is both *hund* and *hundteontig*. Notice also that the lemma for the number 80 (*lxxx*) is *hundeahtatig*, cf. 91078. This particular word is frequently misunderstood and read as ‘one hundred and eighty’.

102698 (e.g.): *Uton* ‘let us’ has the lemma *witon* (cf. B&T) and is analyzed as a 3rd person plural subjunctive. See, however, van Bergen (2013).

102586: *Hiwan* ‘household’ is a masc. pl. and has the masc pl. lemma form (no singular form, according to B&T)

102467 (e.g.): The feminine noun *menigu* ‘crowd’ is indeclinable. We have given it case according to its function, if unambiguous, and taken the number to be singular.

102724: The word for ‘basket’ can be both masculine *wiliga* and feminine *wilige*. In 102724, we have taken it as masculine because of the masculine declension of *fulle*. The lemma is thus *wiliga*.

119074 (e.g.): The lemma for the adverb and subjunction *ponne* is in this corpus *ponne*, even though B&T has *panne* as the main form. That is because the word hardly ever occurs as *panne*!

107844, 107821: The Latin conjunction *et* is used in the text (*Ælfric LS*) instead of *and*. If it is tagged in the morphology as ‘foreign word’, the application will not let us ‘use’ it in the syntax layer. Therefore we have tagged it as ‘conjunction’ and given it the lemma *et*.

107848: *God* ‘the Christian God’ (proper noun) and *god* ‘heathen god’ (common noun) are two different lemmas.
Crist ‘Christ’ (proper noun) and crist ‘false christ’ (common noun) are two different lemmas.

There are two lemmas for the verb answer: andswrian and andwerdan.

There are two lemmas for the verb show, display, appear: ætywan and æteowian.

dgef ‘false christ’ (common noun) are two different lemmas.

There are two lemmas for the verb answer: andswrian and andwerdan.

There are two lemmas for the verb show, display, appear: ætywan and æteowian.

gebringan and gebrengan seem to mean the same thing, ‘bring’, but gebrengan is a weak verb (gebrohte, gebroht), and gebringan is a strong verb (gebrang, gebrungen). We have used the verb forms and the vowels to determine whether it is one or the other. In the corpus examples, the present tense gebringð occurs, which means that the lemma is gebringan, whereas the participle gebroht shows that the lemma must be gebrengan.

It happens, though rarely, that what is represented as one word in the OE text has to be split into two lemmas. In 102475, the word toegaderod is split into to (preposition) and gegaderian (verb), as there is no verb toegaderian, according to the dictionaries. In 102482, upeode was split into up (adverb) and gan (verb), as there is no verb upgan.

ªider. We take this as a non-inflecting adverb. YCOE has it as a dative-marked adverb, presumably because it is an absolute demonstrative (see B&T). We do not mark case on adverbs in our system, however.

Eastron is actually the dative plural of easter, but we have tagged it as non-inflecting, since that is the form found in almost all contexts, since it is used as a subject here, and since Latin and Greek pascha are non-inflecting.

The noun sawel ‘soul’ is feminine, but in this sentence it seems to be used as a masculine or neuter noun, cf. unrot and min.

Usually, we annotate the token þyþi as the instrumental case of the demonstrative determiner se. In this sentence, however, þy functions as adverb in the first case, and subjunction in the second, and they had to be annotated as such in order for the syntax to be correct. The lemma form is þy in both cases.

The lemma for the conjunction ægðer is æghwæðer.

There are two possible words for ‘depart, go’, namely faran (strong) and feran (weak). It is usually not difficult to distinguish between them, since the past tense sg of faran is for (with vowel change), and the past tense sg of feran is ferde (inflection). However, in Ælfric’s Lives of Saints we come across forms...
such as *fieren* (infinitive), *færde* and *færdon*. We have lemmatized these as *feran* as well, due to the inflected past tense.

103603: The word ‘depart’, ‘die’ has two possible lemmas: *forþfaran* and *forþferan*. The most common form is the past tense *forþferde*, which gets the lemma *forþferan* (weak).

107928: There is a plural masculine lemma *burhware* ‘the inhabitants of a city, citizens’ and a singular feminine lemma *burhwaru* ‘the body of citizens’. Both are found in the corpus: the former in e.g. 107928, 91102, 91154, and the latter in 102333.

91163: *bringan* and *bren gan* mean the same (‘bring, adduce, lead, produce, bear, carry’), but are two different lemmas.

### 2.3 LATIN CASE ENDINGS

Proper names quite often get Latin case in texts that are translated from Latin. If the case is unambiguous, we tag it, e.g. in *pa heo becom to Apollonio* ‘When she came to Apollonius’ (118918), where *Apollonio* is the second declension dative/ablative case for masculine nouns in *-us*. The scribes were, however, not consistent in using the Latin case endings: in *Da het se cyng sillan Apollonige pa hearpan* ‘Then the king bade the harp be given to Apollonius’ (118961), *Apollonige* has the OE dative *-e* ending. If it is not possible to determine the case, it is tagged as ‘unspecified’, e.g. in *Lareow Apolloni, ic gife þe*... ‘Master Apollonius, I give you...’ (118976), where *Apolloni* is not the vocative case in Latin.

A list of all occurrences of names with Latin case endings has been made, with their token numbers. This information will be added to the corpus later, so that it will be possible to search for all the names with Latin endings.
3 SYNTAX

3.1 SENTENCE DIVISION
The system for sentence division is as follows: Coordinated clauses are split into individual sentences if there is an overt subject. If the subject is ellipted, the sentence is not split. In other words, a sentence like *He came and he saw and he conquered* will be split into three separate sentences: *He came* / *and he saw* / *and he conquered*, whereas a sentence like *He came, saw and conquered* will be kept as one sentence with coordinated PREDs and slashes (secondary dependencies) from the verbs *saw* and *conquered* back to the subject *he*. If the sentence is extremely long, we may have to split it for practical reasons, so the tree does not become too large.

The Romance languages Old Spanish and Old Portuguese are null-subject languages, so for those languages it was decided to simply split the sentences wherever there is a finite verb, unless coordinated verbs share arguments. Old French is different from Old Spanish and Old Portuguese on the one hand, in that it has fewer null subjects than OS and OP, and different from Old English on the other hand, in that it is in fact a pro-drop language, so there was discussion concerning how to deal with sentence division in that language. In the end we decided to use the same system in OE and OF.

Sometimes our system seems to be at odds with either the original manuscript’s or the edited version’s sentence division. For example, in (12):

(12) *Eugenia ða wunode on þam mynstre mid wærlicum mode þeah þe heo mæden wære, mid hyre twam cnihtum, uncuð gehwam, And heold on hyre þeawum halige drohtnumge...* (107842)

Eugenia then dwelled in the minster with manly mind though *she* she maiden were, with her two knights, unknown everyone-DAT, And observed in her conduct holy service...

‘Eugenia then dwelled in the minster with a man’s mind, though she were a maid, with her two servants, unknown to every one, and observed in her conduct the holy service...’

there is a capital letter in *And*, preceded by a comma. We have merged these to one sentence, since the second clause does not have an overt subject. The capital letter indicates that there is a major division there, whereas the comma indicates that there is not...

In general, we follow the sentence division of the edited version in cases of doubt (our aim is not to do manuscript philology). For example, 102555 *Se hæfde on byrgenum scref* ‘he/who had his dwelling among the tombs’ can either be interpreted as a se-relative clause (not all subclauses have a clause-late verb) or as a main clause. However, the editor has clearly interpreted it as a main clause, and we trust the editor’s judgement in this case.

If there are two verbs that seem to be coordinated, but there is no coordinating conjunction between them, we split. An example is (13), where we split between *ofsceamod* and *wende*.
(13)  *Đa wenđ Melantia micclum ofsceamod, wende þæt heo wolde hyre word ameldian...*
(107978-107879)
then became Melantia much ashamed, supposed that she [Eugenia] would her conversation betray...
‘Then was Melantia greatly ashamed, and supposed that she would betray her conversation...’

3.2  VERB PHRASES WITH ‘MODAL AUXILIARY’ + INFINITIVE
Complex verb phrases with a ‘modal auxiliary’ + infinitive are analyzed as PRED + XOBJ, as in (14). This analysis gives a biclausal structure, rather than a flat structure for the sentence, and this analysis was chosen because the auxiliaries were not fully grammaticalized in Old English; they retained some of their lexical force (Traugott, 1992.). The alternative analysis is to take the lexical verb as the head and the auxiliary as an AUX element, and this is what the Menotec project chose to do for Old Norse, cf. (15), where *gifta* is the head and *villdi* AUX.

Since the OE structure is biclausal, the annotator has to decide which adverbials and arguments belong with which verb. Adverbials of time and place + the negation are usually dependent on the uppermost verb, whereas arguments are usually dependent on the lowest (lexical) verb, e.g. *ealle bigspell* in (14).

As the tree in (14) also shows, our analysis allows us to show that *ge* is the subject of both the auxiliary *mage* and the infinitive *witan*, through the XOBJ function. The label XOBJ is used for open predications that do not supply their own subject, but get their subject via coreference relations in the sentence (cf. PROIEL guidelines 10.9). The slash (stipled line) from *witan* to *ge* shows that the subject of *witan* is *ge* (XSUB means ‘external subject').

(14)  *& hu mage ge ealle bigspell witan?* (102499)
and how may you all parables know
‘and how may you know all parables?’
(15) *En faðer hennar villdi engvm kosti gifta hana* (Old Norse, Menotec 14080)

but father her would no-one marriage.DAT give her

‘But her father would not give her to anyone in marriage’

The only instance of a complex verb phrase in which the lexical verb gets the function PRED in OE is in phrases consisting of the verb *habban* ‘have’ + participle. Hence, in 119086, *Þu hafast gecoren þone wer þe me wel licað* ‘You have chosen the man that is pleasing to me’, *gecoren* is PRED and *hafast* is an AUX dependent on the PRED, i.e. the same structure as in the Old Norse example in (15).

The verb *onginnan/beginnan* + infinitive also gets the XOBJ analysis, e.g. 102587.

### 3.3 ACCUSATIVE WITH INFINITIVE

Accusative with infinitive (AcI) occurs especially after verbs of the senses, as in (16). In this kind of sentence, the accusative + infinitive unit is regarded as an argument of the finite verb, and the construction is thus analyzed as COMP. The subject, *ænne nacodne cnapan* in (16), gets its semantic role (theta-role) from the non-finite verb; it is therefore dependent on the infinitive, and gets the function SUB.

(16) *...þa geseah he ænne nacodne cnapan geond þa stræte yrnan* (118863)

...then saw he a naked boy through the street run

‘...then he saw a naked boy run through the street’
(17) is an example of an infinitive without an accusative:

(17)  *Ich gehirde secgan þæt ic wære fordemed* (118797)
     I heard say that I was condemned
     ‘I heard say that I was condemned’

In this case, *secgan þæt ic wære fordemed* gets the function COMP, and the infinitive does not have a subject.

It also happens that a perception verb is followed by an accusative + a present participle instead of an infinitive, as in (18).

(18)  *& he geseah hi on rewette swincende* (102730)
     and he saw them in rowing toiling
     ‘and he saw them toiling in rowing’
In such cases, we choose an XOBJ analysis, not a COMP analysis, although the meaning is very similar: *He saw them toil* and *He saw them toiling* differ only in the durative meaning of the latter. However, infinitives and gerunds are different syntactically, which is why the accusative is analyzed as an argument of the main verb when it is followed by a present participle. In other words, *hi* is OBJ to *geseah* (rather than subject to *swincende*) and *swincende* is XOBJ to *geseah*, with *hi* as XSUB (external subject).

In (19), we have an example of an accusative with infinitive with an impersonal verb.

(19)  
\& on ealle þeoda ærest gebyrdæ beon þæt godspel gebodud (103315)

and among all nations first behoves be that-DEM gospel published 
‘and it is fitting that the gospel should first be published among all nations’

3.4 **HATAN** + INFINITIVE

A slightly problematic feature of OE is that both perception verbs and verbs of causing and commanding are followed by accusative + infinitive, unlike Old Norse, in which commanding verbs, e.g. *bjóða* ‘bid (someone to do something)’, are followed by a dative object. In OE, however, the commanding construction is very similar to the accusative with infinitive construction, so it is difficult to know which ones are true AcI constructions in which the AcI as a unit is dependent on the finite verb, and which ones are constructions in which the
accusative and the infinitive are separate arguments of the finite verb, as also repeatedly commented on by Mitchell (1985 II), who discusses these constructions extensively. (Cf. also PROIEL guidelines, section 10.6.)

We decided, after much deliberation, to analyze Old English in the same way as Old Norse, with perception verbs taking AcI, and thus COMP, as shown in 3.3, and causative verbs and verbs of commanding taking an XOBJ construction; in other words, the accusative element is in the latter case analyzed as an argument of the finite verb. The motivation for this is the existence of various constructions for a verb such as *hatan* ‘to order, command, bid’, such as the ones exemplified in (20) and (21). Especially (21) suggests that the accusative element is in fact an argument of the finite verb.

(20)  *and het [his hyredmen ealle] [him aweg gan] (118656)*  
and bade his domestics all him-DAT away go  
‘and bade his domestics all go away from him’

The analysis is thus not *het [his hyredmen ealle him aweg gan]*.

(21)  *þa het ic [eallne þone here] [þæt he to swæsendum sæte]*  
(Letter of Alexander the Great to Aristotle, 10th-11th c., from Mitchell 1985 II: 871)  
then bid I all-ACC the-ACC army-ACC that it to meal sit-SBJV  
‘then I bid the entire army to be seated for the meal’

Example (22) is an interesting example of a coordinated structure in which *hine* in the first clause is the object of *gescrudan* – the king ordered (an implied someone) to clothe him (i.e. Apollonius) – and the second, coordinated, clause is a normal *hatan* structure, where *hine* is analyzed as an argument of the main verb – the king commanded him (i.e. Apollonius) to go in – and also the subject of the infinitive.

(22)  *Ða het se cyngc hine sona gescridan mid wurðfullan scrude and het hine in gan to ðam gereorde.* (118902)  
then commanded the king him immediately clothe-INF with honourable clothing and commanded him in go-INF to the repast  
‘Then the king commanded him to be clothed immediately with honourable clothing, and bade him enter to the repast’
Notice the secondary dependencies – the slashes: In the first clause the slash goes from the infinitive *gescridan* to *het*, because there is no explicit external subject for the infinitive. In the second clause, the slash goes from the infinitive *gan* to the object of *het, hine*, which is also the subject of the infinitive.

Example (23) shows yet another type of *hatan* construction. Here, the external subject of the infinitive, *geclipian*, is the same as the subject of *het*, namely Apollonius, and the slash therefore goes back to *Apollonius*. Apollonius both bids and calls.

(23)  *and Apollonius het hine eft to him geclipian and cwæð to him: (118787)*

and Apollonius bade him again to him call-INF and said to him
‘and Apollonius bade him again be called to him and said to him
Sentence (24) is an example containing the commanding verb *biddan*:

(24) *and þa men ealle arison and gretton þone cyngc and ða cwene and bædon hig gesunde beon and ham gewændon.* (118982)
and the men all arose and greeted the king and the queen and bade them healthy be 
and home went
‘and the men all arose, and greeted the king and the queen, and bade them farewell, 
and went home’

*Biddan* is analyzed like *hatan*; i.e. *hig* is regarded as belonging in the argument structure of *biddan*, and gets the function OBJ to *bædon*, and in addition there is a slash from *beon* to *hig*, showing that *hig* is the external subject of *beon*. 
See Bech & Eide 2011 for a discussion of the problems of AcI and commanding verbs in Romance.

3.5 **VERB PHRASES WITH BE + PAST PARTICIPLE**

Passive constructions, e.g. (25) and (26), which consist of *be* + past participle have also been analyzed as XOBJ constructions. This analysis means that the relation between participles and adjectives is emphasized. The participle is analyzed in exactly the same way as a nominal or adjectival subject complement would be analyzed (and the participle may carry case inflections). So in a sentence like *He is ugly*, *ugly* would be analyzed as an XOBJ to *is*, with an XSUB slash back to the subject *he*, just like the participle *todaeled* is analyzed in (25). It is a matter of discussion whether an adjective can have a subject, but we will not enter into that discussion here. The disadvantage of this analysis for the passives is that it will be very difficult to search for passives, except those (few) passive constructions that have an agent phrase (tagged AG), cf. (26).

(25) & gif Satanas winð ongen hine sylfne he bidó todaeled (102458)
and if Satan fights against him self he is divided
‘and if Satan rises up against himself, he is divided’
(26) *Da ḷa ḷæt maedæn gehirde ḷæt hire was alyfed fram hire ūæder ḷæt heo ær hyre sîlf gedon wolde, ḷa cwæð heo to Apollonio* (118942)

when when the maiden heard that her-DAT was allowed from her father that she before her self do would, then said she to Apollonius

‘When the maiden heard that that was allowed her from her father what she herself wished to do, she said to Apollonius’
Example (27) is an instance of the ‘was called’ passive, where the participle *haten* is analyzed as XOBJ to the main verb *wæs*, and the subject complement is analyzed as XOBJ to *haten*, with a slash back to *haten*, which again slashes to the subject *seo*.

(27) ..... *seo wæs haten Ciðnus* (116674)

... she was called Cydnus

‘... it [it is a river] was called Cydnuz’

In the PROIEL project, passives have been analyzed in this way (PRED + XOBJ) for all their languages (i.e. Gothic, Greek, Church Slavonic and Armenian), except Latin, where passives are regarded as more grammaticalized than in the other languages. This is because this construction is regarded as a form in Latin grammars and is part of the verbal paradigm.²

3.6 VERB PHRASES WITH *BE* + PRESENT PARTICIPLE

As regards *be* + present participle, we distinguish between two constructions: 1) sentences in which the participle is part of the verb phrase (resembling the modern English progressive, but with a biclausal structure) and 2) sentences in which *be* is a position or presentational verb, and the present participle has an adverbial function. In the former case, the present participle gets an XOBJ analysis, whereas in the latter case, the present participle gets an

² Dag Haug (pers.comm.) recognizes that this a rather subjective judgement, and says that if they were to do it again, they would probably have used the AUX-PRED analysis for some of the other languages as well.
XADV analysis. Examples (28) and (29) are examples of the former. In (28), \( \text{\textit{wæs}} \) is PRED and \( \text{\textit{bodigende}} \) and \( \text{\textit{adrifende}} \) are XOBJ, and in (29), \( \text{\textit{wæs}} \) is PRED and \( \text{\textit{sprecende}} \) is XOBJ.

(28) \& he \textit{wæs bodigende on heora gesamnungum \& ealre Galilea, \& deofolseocnessa ut adrifende} (102343, see also PROIEL 10354 for the corresponding Latin sentence) and he was preaching in their synagogues and all-DAT Galilee, and devil-sickness out driving
‘and he was preaching in their synagogues throughout Galilee, and casting out devils’
(29)  *Mid þi þe he þas þingc wæs sprecende to him sīlfum, þa færinga geseah he sumne fiscere gan* (118844)

with that *he* [complex subjunction ‘while’] he these things was speaking to him self, then suddenly saw he a fisherman go

‘While he was speaking these things to himself, he suddenly saw a fisherman going’
In (30), on the other hand, the present participles are analyzed as having an adverbial function: *There were some of the scribes, sitting and thinking.*

(30) *Par wæron sume of ðam bocerum sittende, & on heora heortum þencende.* (102364, see also PROIEL 10373 for the corresponding Latin sentence)
there were some of the scribes sitting and in their hearts considering
‘some of the scribes were there, sitting and considering in their hearts’

Example (31) is similar to (30), but notice that the verb in (30) is analyzed as a position verb, and *par* as an XOBJ, whereas the verb in (31) is analyzed as a presentational verb, with *par* as an EXPL(ative) and *embe þone munt* as an ADV. The reason for the difference is that *sume of ðam bocerum* in (30) is specific, so it cannot be a presentational sentence. See also section 3.7 for further comments on the difference between location sentences and presentational sentences, and section 3.18 for further comments on existential *there*.

(31) *Par wæs embe þone munt mycel swyna heord læsgende* (102570)
there was about the mountain great swine-GEN herd feeding
‘There was a great herd of swine feeding near the mountain’
Another example of the XADV analysis is (32), where *on westene* is XOBJ (see 3.7) and the present participles are analyzed as XADV, i.e. adverbials with an external subject (*Iohannes*).

(32)  *Iohannes wæs on westene fulligende & bodiende dædbote fulwiht on synna forgyfenesse* (102277)
      John was in wilderness baptizing and preaching repentance-GEN baptism for sins-GEN forgiveness
      ‘John was in the wilderness, baptizing and preaching the baptism of repentance for the forgiveness of sins’
If there is another verb than *be*, of the type *He came running*, the present participle is also an XADV. An example is (33) (see also e.g. 107857 in the corpus).

(33) & se unclæna gast hine slitende & mycelre stefne clypiende him of eode (102321) and the unclean spirit him tearing and loud-DAT voice-DAT crying him from went ‘and the unclean spirit, tearing him and crying with a loud voice, went from him’
3.7 ADVERBS AND PREPOSITION PHRASES WITH THE VERB BE

The verb *be* can be a copula, part of a passive verb phrase, a position verb, or a presentational (existential) verb. The type of *be* has consequences for our analysis of adverbs and preposition phrases (PP) in the sentence. If *be* is a copula, the subject complement (whether it is an adjective or a PP) gets the function XOBJ, and additional adverbs or PPs get the function ADV (copulas do not take adverbial arguments (OBL)). The same analysis is used for passives, except that such clauses may also have adverbial arguments, which are analyzed as OBL (*The glass was placed on the table*), and/or a PP that functions as AG(ent) (*The glass was placed on the table by my mother*).

However, what needs to be especially kept in mind by the annotators is the analysis of adverbs and PPs in sentences with position verbs and presentational verbs. With position verbs, e.g. *He was in Galilee*, the PP gets the function XOBJ, which means that it is regarded as an argument. With presentational verbs, e.g. *There was a man in Galilee*, the PP gets the function ADV, i.e., it is not an argument. Examples can also be found in (31) and (32): In (31), *embe þone munt* is an ADV (presentational sentence), and in (32), *on westene* is an XOBJ (position verb). It may sometimes be difficult to determine whether *be* is a position verb or a presentational verb. For example, in 90688 *& senop was æt Clofeshoo* ‘and synod was at Cliff’s-Hoo’, is the meaning ‘a synod was at Cliff’s-Hoo’ or ‘there was a synod at Cliff’s-Hoo’? In the corpus, it has been analyzed as a presentational sentence, with *æt Clofeshoo* as an ADV.

With other position verbs, such as *sittan* ‘sit, remain, dwell’, *licgan* ‘lie’, *standan* ‘stand’, the PP or adverb gets the function OBL.
This distinction is also made by the PROIEL project for their languages. The tree bank for Modern Norwegian would also regard *in Galilee* in *He was in Galilee* as a predicative element. The Menotec project, on the other hand, analyzes such elements as OBL, i.e. as obligatory adverbial elements, not predicatives. Both analyses have their strong and weak points. The disadvantage of the ISWOC/PROIEL analysis is that no difference is made between sentences such as *She was happy* and *She was in London*, and it may be argued that there is a grammatical difference between them (see Hasselgård 2010: 14-15). The disadvantage of the Menotec scheme is that it is sometimes hard to judge when a verb is a position verb and when it is not.

### 3.8 VERB PHRASES WITH *HABBAN* + PARTICIPLE

This is the only verb phrase in which the auxiliary is tagged as AUX and the main verb as PRED, i.e. it is regarded as a flat, monoclusal structure. (See also comments at the end of 3.2.) An example is (34).

(34) *...þu hafast nu geedniwod his ealde sar* (118940)

...you have now renewed his old grief

‘...you have now renewed his old grief’

It is uncertain whether this type of verb phrase at this stage was truly verbal, or whether the participle still had some kind of adjectival function, since some participles in *habban* constructions get case endings still. In other words, an alternative analysis would be PRED + XOBJ, but we have decided to analyze all these verb phrases in the same way, regardless of whether the participle is inflected or not.

---

3 Cf. e.g. their sentences 10310 and 10582 vs. 10326.
3.9 SUBJECT COMPLEMENTS

See comments on subject complements in 3.5 and 3.7. In addition, it should be mentioned that we analyze constructions like *What is this* with *what* as XOBJ. It is sometimes difficult to know what is the subject and what is the complement, but we find it reasonable to assume that it is the question word that represents the predicated quality/property/identity. (35) is an example.

(35) *Hwæt* is þeos niwe lar... (102324)

What is this new doctrine...
‘What is this new doctrine...’

Of course, this does not mean that *hwæt* always functions as XOBJ. In (36), *hwæt* functions as SUB(ject) and *þæes* is the XOBJ, i.e. the subject complement.

(36) *ac he nyste hwæt þæes sopes wæs, for þæm he hit self ne geseah* (115590)

but he not-knew what that-GEN truth-GEN was, for that he it self not saw
‘but he did not know what of that was true, because he did not see if for himself’
3.10 RELATIVE CLAUSES

Old English relative clauses are constructed in five different ways: with the indeclinable relative particle *pe* (cf. the Old Norse relative particle *er*), with the demonstrative *se* as the relative pronoun, with a combination of the demonstrative and the relative particle: *se pe*, of which there are two subtypes, or as headless relatives. These possibilities are reflected in the syntactic analysis.

3.10.1 *pe* relatives

The relative particle *pe* is analyzed as a subjunction, and it is thus the head of the relative clause, and not an argument (relative pronouns are arguments). This analysis can be seen in (37) and (38), where *pe* is the head of the relative clause, and the verb is dependent on *pe*. Notice that in (37), the relative clause has the label APOS, which means that it is non-restrictive (appositional), whereas in (38) it is labelled ATR, i.e. restrictive (attributive).  

In other words, unlike other corpora (we think), we distinguish between restrictive and non-restrictive relative clauses, and it is usually not very difficult to make this distinction. If the annotator is unable to determine whether the relative clause is restrictive or non-restrictive, it is possible to use the supertag REL, which just means ‘relative clause’.

(37)  *se Iohannes pe ic beheafdode se aras of deaðe* (102668)

the John that I beheaded, he arose from death

‘It is John, whom I beheaded; he is risen from the dead’

(38)  *& to him com mycel menegeo ymbe Tirum & Sidone gehyrende þa þing þe he worhte* (102435)

I actually have the impression that most *pe* relatives are restrictive clauses, whereas *se* relatives are mostly non-restrictive, but I have not carried out a systematic search to check this. See table in Mitchell 1985 II:175, for OE poetry.)
and to him came great crowd around Tyre and Sidon hearing the things that he did
‘and a great crowd, hearing about the great things that he did, came to him from
around Tyre and Sidon’

(37)

(38)

For further examples of *he* relatives, see e.g. 118967 (ATR), 119081 (ATR), 118978 (ATR),
118989 (APOS), 118900 (ATR).

### 3.10.2 *se* relatives
The second type of relative clause is the type in which the clause is introduced by a form of
the demonstrative determiner *se*. In the morphology, these are tagged as demonstrative
pronouns, so we do not have a separate category for relative pronouns. The idea is that the
difference between demonstrative pronouns and relative pronouns is not a morphological one,
but a syntactic one, and it is shown in the syntax.

In (39), the head of the (appositive) relative clause is the verb, which gets the tag
APOS, and the demonstrative (nominative plural) *pa* functions as SUB in the relative clause.
It gets case according to its function in the relative clause.5

(39)  *pa æt nyhstan comon ðar gan ongean hy þry gelærede weras and æpelborene, pa langle ær girndon þæs cyninges dohtor* (119009)

5 What I have taken as a relative clause here *might* also be analyzed as a main clause ‘they had long before
desired the king’s daughter’.
then at length came there go-INF towards them three learned men and noble, who long before desired the king’s daughter

‘Then, at length, there came walking towards them three learned and noble men, who long before had desired the king’s daughter’

Sometimes the demonstrative pronoun occurs in a prepositional phrase, and thus functions as a prepositional complement that gets case from the preposition, cf. (40), in which the PP be þære functions as OBL (an obligatory adverbial) in the relative clause. Notice also, by the way, the þe relative with an APOS function at the end of the sentence, in which the subjunction þe cannot be analyzed as a prepositional complement to the preposition embe. Embe is here analyzed as an OBL by itself (cf. also 115601).

(40)  His wif wæs gecyged Claudia, be þære he gestrynde twægen suna, Auitum and Særgium, and ane dohtor, Eugenian, þe we embe spærcap (107810)

his wife was called Claudia, by whom he begot two sons, Avitus and Sergius, and one daughter, Eugenia, þe we about talked

‘His wife was called Claudia, by whom he had two sons, Avitus and Sergius, and one daughter, about whom we have talked’
Relative clauses of the type *The place where he lived* are analyzed in the same way as *se* relatives, except that in such clauses, the ‘relative adverb’ gets the function ADV or OBL in the relative clause. Note that we do not have a morphological category ‘relative adverb’; *þær* is tagged as a general adverb (see 2.1.19). See 118746 for an example.

### 3.10.3 *se þe* relatives

*Se þe* relatives are constructed with a combination of the demonstrative *se* and the relative particle *þe*. There are two subtypes of *se þe* relatives. (See Haugland 2007: 307-309 for a nice overview.) In the first type, type 1, *se* is head, taking its case from the main clause, and *þe* is a dependent of *se*, taking the relative clause verb as its dependent (*þe* is thus a subjunction, as in *þe* relatives). The meaning is ‘the one who/which’. In such clauses, the demonstrative determiner functions as an APOS to the antecedent, and the *þe* clause is attached to the demonstrative as an ATR. For example, in (41), the second *se* is the subject of the main verbs.
naefð and bið. The first se attaches to the second se as APOS, and þe attaches to the second se as ATR.

(41) **se þe ðone halgan gast bysmerað, se naefð on ecnysse forgyfenesse, ac bið eces gyltes scyldig** (102463)

he þe the holy ghost mocks, he not-has in eternity forgiveness, but is perpetual-GEN sin-GEN guilty
‘he who blasphemes against the Holy Ghost shall never have forgiveness, but is in danger of eternal damnation’

A similar example to (41) is 102473, while in 107980 ...þæt ða cristenan tocneowan þæt Crist wæs mid hyre, se ðe hwilon ær þone halgan Petrum be ðære handa gelædde... ‘so that the Christians might know that Christ was with her, he who led the holy Peter by the hand...., se is APOS to Crist, and ðe is attached to se as ATR.
Sentence (42) contains both a *se* relative, and a *se he* relative of type 1, both functioning as APOS.

(42)  *Þa geseah hine sum his cuðra manna se wæs Hellanicus genemnod, se he ærest þider com* (118763)

then saw him one his known-GEN men-GEN who was Hellanicus called, he *he* first thither came
‘Then one of his men who was called Hellanicus, who had first come thither, saw him’

Other examples of *se he* relatives type 1: 102889, 102890.

In the second type of *se he* relative, type 2, the demonstrative gets case from its function in the relative clause rather than from the antecedent, and *he* is merely a reinforcing particle. The meaning is ‘who/which’ rather than ‘the one who/which’. Since there are no unambiguous cases in the corpus, (43) is taken from Haugland (2007: 308, her example (8.39), from ÆCHom). In (43), *seo* is in the nominative case, whereas its antecedent *ðære sunnan* is in the dative case.

(43)  *Oðer is se leoma oððe beorhtnys. æfre of ðære sunnan seo ðe onliht ealne middangeard*

second is the beam or brightness. ever from the-DAT sun-DAT she-NOM ðe illuminates all earth
'The second is the everlasting beam or brightness of the sun, which illuminates all the earth'

In clauses like this, the analysis will be similar to that of *se* relatives; i.e. the head of the relative clause is the verb (rather than *þe*), which gets either the function APOS or ATR, and *se* is analyzed according to its function in the relative clause: in (43), *seo* is SUB(ject). The reinforcing particle *þe* is attached to *seo* as an AUX element. A possible example from the corpus is 102893.

In other words, *se þe* relatives of type 1 are analyzed similarly to *þe* relatives, and *se þe* relatives of type 2 are analyzed like *se* relatives. If my hunch about the difference between *þe* and *se* relatives is correct (and Mitchell’s claim can be extended to prose as well, see footnote 4 above), to the effect that *þe* clauses are usually restrictive and *se* clauses usually non-restrictive, this means that there might also be a difference in meaning between *se þe* relatives type 1 and *se þe* relatives type 2. Type 1 restricts the reference of the antecedent, as in (41), whereas type 2 elaborates on the antecedent, as in (43).

As mentioned above, there are no unambiguous cases of the second type of *se þe* relative in the corpus (so far). In order for such a clause to be unambiguous, the relativized noun phrase must be a prepositional complement, and the case requirement must be different for the antecedent and the demonstrative. If, for example, both the antecedent and the relative determiner require nominative case (which is usually the case), it is not possible to tell whether it is a *se þe* relative of type one or type two. In such cases, the analysis for type 1 is used by default.

One such ambiguous case is found in the second relative clause in (42) above: *se þe ærest þider com*. The meaning is either ‘he who came there first’, or ‘who came there first’, but we cannot tell which, since *Hellanicus* is also a nominative element. I think this relative clause is most likely a *se þe* relative of type 2, since it does not restrict the reference of the antecedent, but note that *we nevertheless analyze all such structurally ambiguous cases in the same way as se þe relatives of type 1*.

### 3.10.4 *eall þæt* constructions

Some sentences are of the type in (44):

(44) *þa foron hie... & genamon eal þæt þær binnan wæs* (91109)

*then went they... and took all that there inside was*

*‘then they went... and took all that was inside there’*

*Eal* is accusative and *þæt* must be nominative, as it is the subject of *wæs*. In other words, *þæt þær binnan wæs* is analyzed as an ATR (restrictive) relative clause to *eal*, with *þæt* as the subject of *wæs*. It seems a bit counterintuitive to do it like this – intuitively *þæt* feels closely connected to *eal*, and *þæt* is also an accusative form. Could it be that this type of construction is like the *þe* relatives, so that *þæt* is really not a relative pronoun in these constructions? See also 91144.
3.10.5 Headless (nominal) relatives
Sometimes there is no antecedent for the relative clause, so that the relative clause has a nominal rather than an adjectival function. In such cases, the verb of the relative clause is the head, and it is tagged according to the sentence function of the relative clause, usually SUB or OBJ. In (45), the nominal relative clause functions as subject (SUB), in (46) it is a subject complement (XOBJ), and in (47), it is an object (OBJ).

(45) and nis naht *pæt þu sægst* (118720)
    and not-is naught that you say
    ‘and what you say is naught’

(46) *sy pæt þu sy* (118847)
    be that you are
    ‘be what(ever) you are’

(47) *Nim pæt ic þe to sillenne habbe and ga into ðare ceastre* (118854)
    take that I you to give have and go into the city
    ‘take what I have to give you, and go into the city’
Further examples of headless relative clauses can be found in 118942, 119047, 119061, 118732, 102352, 102681 (swa hwæt swa), 118958.

Sometimes the headless relative clause is dislocated and picked up by a resumptive pronoun in the main clause, i.e. sentences of the type Whoever solves this riddle, let him get my daughter. In such cases, the headless relative clause is analyzed as an APOS to the resumptive pronoun, as in (48).

(48) And gif Ȱu wille þisum þingum gehyrsum beon,  ic swerige Ȱe þurh mines rices mægna þæt swa hwæt swa Ȱu on sæ forlure ic Ȱe þæt on lande gestaðelige (119005)
and if you will these-DAT things-DAT obedient be, I swear you through my-GEN kingdom-GEN powers that so what so you at sea lost I you that on land restore
‘And if you will be obedient in these things, I swear to you by the powers of my kingdom, that whatever you have lost at sea, I will make it good to you on land’
Further examples of headless relative clauses with this function are found in 118678, 118679, 118755, 118756, 118783, 18784.

Finally, I would like to mention two examples of special relative constructions. The first one, in (49), involves the construction wið þæm þe, meaning ‘with which ðe he persuaded the Greeks to withstand Alexander’. If the construction had just been wið þæm, the analysis would have been straightforward: wið þæm would have been ADV in the relative clause. However, the þe complicates matters, since it is analyzed as a subjunction in regular relative clauses. The chosen analysis is shown below, where þe is the head of the relative clause, and wið þæm is analyzed as an OBL to the verb gelærde.

(49) Þæt wearð ærest from Persum, þa hie sealdon Demostanase þæm philosophe licgende feoh wið þæm þe he gelærde ealle Crecas þat hie Alexandre wiðsocen (116659)
that happened first concerning Persians, when they gave Demosthenes the philosopher ready money with which þe he persuaded all Greeks that they Alexander withstood ‘It first happened that the Persians gave Demosthenes, the philosopher, ready money, with which he seduced all the Greeks to strive against Alexander’
The second special case is a relative clause without a verb. In (50), an empty verb and an empty coordinated conjunction have been inserted, and the subject of the relative clause is implied. The meaning is thus: ‘Pity me, who is naked and shipwrecked’.

(50) \textit{gemildsa me nacodum, forlidenum} (118848)  
\textit{pity me naked, shipwrecked}  
‘pity me, naked, shipwrecked!’
3.11 COMPARISON

We analyze comparisons in the same way as the Menotec project (see Menotec guidelines, section 10.6), and unlike the PROIEL project; i.e. we analyze comparisons of the type shown in (51) to have an unexpressed verb, whereas PROIEL regards them as ‘reduced’ sentences. In (51), for example, *wælhreowra* is the subject complement, analyzed as XOBJ. We attach the subjunction *þonne* to the subject complement as COMP, i.e. a complement clause which is dependent on the adjective, and insert an empty verb, to which *Antiochus se cyngc* is SUB(ject). In other words, the underlying sentence would be ‘you are crueler than king Antiochus is’. The PROIEL project, which does not operate with empty verbs in these constructions, would analyze *þonne* as an OBL to *wælhreowra* and *Antiochus se cyngc* as a to *þonne*.

(51) *Eala þu sæ Neptune, manna bereafigend and unsceðigra beswicend, þu eart wælhreowra þonne Antiochus se cyngc* (118839)
alas you see Neptune, men-GEN bereaver and innocents-GEN deceiver, you are crueler than Antiochus the king
‘Alas, Neptune of the sea, bereaver of men and deceiver of the innocent, you are crueler than Antiochus the king’
Sometimes comparisons actually have an expressed verb, as in (52):

(52) ic me næfre bet ne baðode þonne ic dide todæg (118885)
I me never bathed not better than I did today
‘I never bathed better than I did today’.
For further examples of this type of comparison, see 115595 (micle læssa þonne), 115596 (lengra ðonne), 116668 (swiþor þonne), 115607 (ma ðonne), 116670 (na ma þonne), 102986 (betere þe is þæt... þonne...)

Another type of comparison is found in constructions of the type swa...swa ‘as...as’, and swa...þæt ‘so...that’, and as shown in (53) and (54), we analyze these as COMP as well. In other words, the þæt clause or swa clause is a COMP to the first swa, the ‘so’ element. This is different from the Menotec scheme for this type of clause; they would give these clauses the function APOS (cf. Menotec 36100). However, we have decided to treat all the constructions involving comparison in the same way.

(53)  *þa wæs he swa feor norþ swa þa hwælhuntan firrest faraþ* (115573)  
then was he as far north as the whalehunters furthest go  
‘Then was he as far north as whale hunters ever go’

(54)  *Hwa wæs æfre swa dirstiges modes þæt dorste cynges dohtor gewæmman ær ðam dæge hyre brydgifta and him ne ondrede þæs cyninges ire?* (118667)  
who was ever so daring-GEN mind-GEN that durst king-GEN daughter defile before the day her nuptials-GEN and him not dread the king-GEN ire?
‘Who was ever of so daring mind that dared defile the king’s daughter before the day of her nuptials and not dread the king’s ire?

For further examples of this type of comparison, see 118741 (swa micle lufe þæt), 118682 (swa same swa), 115573 (swa feor norp swa), 102396 (swa lange swa), 102397 (swa lange tide swa), 118988, 107969 (swa raðe swa), 119006 (swa wel swa), 118654 (special construction with to ðam swiðe þæt).
A construction that is related to the comparisons is the one found in (55). Here, the first *swa* is attached to the second *swa* as an AUX element, in the same way as with complex subjunctions. AUX means ‘auxiliary word’ (and not auxiliary *verb*), and is used to mark grammatical words that do not bear a syntactic function towards their heads, cf. PROIEL guidelines, section 7.9. An empty verb is inserted, and *fæder* is an XOBJ to this verb.

(55)  *Da oncneow* Philippus *swa swa fæder Eugenian* (107908)
then recognized Philip as as father Eugenia
‘Then Philip, as a father, recognized Eugenia’

Sentence 107905 is another example of this construction.

In (56), the *swa swa* clause functions as subject complement, and gets the tag COMP rather than XOBJ, since COMP ‘overrules’ other relations whenever an argument is sentence-formed, cf. PROIEL guidelines, section 10.1.1.

(56)  *& he gemiltsode him for þam þe hi weron swa swa sceap þe nanne hyrde nabbað* (102704)
and he pitied them for that *þe* they were as as sheep that no shepherd not-have
‘and he had compassion with them, for they were as sheep not having a shepherd’
Other examples of comparisons: 118682, 118741, 102313, 102666, 102864, 102897, 91081, 91084, 103274, 122477.

3.12 APPOSITIONS

In appositions (APOS), the element order decides which is the head, and which is the apposition. So in *Apollonius, se tirisca*, *Apollonius* is the head and *se tirisca* the apposition. In *ic, Apollonius, se tirisca ealdorman*, *ic* is the head, and *Apollonius* and *se tirisca ealdorman*
are two separate appositions to ic. In hlaford Apolloni, hlaford is the head and Apolloni the apposition.

We also use the APOS analysis with place names, if one is a specification of the other, e.g. The army went into Mercia, to Nottingham (90836). In such cases, the specifying preposition phrase is a dependent on the preposition of the first prepositional phrase. The same is the case when there is an adverb followed by a specifying prepositional phrase, e.g. there, in that place (90839).

3.13 COMPLEX SUBJUNCTIONS

Complex subjunctions can be analyzed in various ways. The Menotec project has adopted a conservative approach whereby complex subjunctions are analyzed according to their morphology. For example, they would analyze OE mid ði þe in the following way: mid is ADV (or APOS if the construction is dislocated), ði is OBL (governed by the preposition), þe is APOS. (Cf. Menotec guidelines, section 10.3). The disadvantage of this analysis is that we get some weird-looking dependency trees, in which the meaning/importance of the complex subjunction seems exaggerated. We have therefore decided to do it in a different way, assuming that the complex subjunctions were grammaticalized to a much greater extent. Thus, in mid ði þe, mid and ði are analyzed as AUX to the subjunction þe. Thus, only þe really has a function. (Ann Taylor agrees with this analysis (personal communication).)

List of complex subjunctions:

- after þam þe: 118726. after and þam are AUX to þe.
- ær þam þe: 118760, 33815. ær and þam are AUX to þe.
- buton þæt: 118680. buton is AUX to þæt.
- forðam þe: 118748. Also forðon þe. Both are subjunctions in the morphology, but forðam is AUX to þe.
- forði þe: 91151. Same as forðam þe.
- for þæm: 115577, 115590, 115605. for is AUX to þæm, which is the subjunction. The lemma has to be þæm, then – same problem as with þi in mid þi (see below). Menotec runs into the same problem, even if their basic analysis is different.
- for þon: 116692. þon is the subjunction, and the lemma is also þon.
- mid þam þe: 119028.
- mid þi: This one is a problem, because þi has to be tagged as a subjunction in the morphology, otherwise there will be problems in the syntax. But in the three-part subjunction mid þi þe, we tag þi as a demonstrative pronoun in the instrumental case (lemma: se), and þe as the subjunction. However, it is difficult to see any other way of doing it here than to take þi as the subjunction and lemmatize it as þi.
- mid þi þe: mid and þi are AUX to þe.
• *œþ þæt*: ‘until’. *œþ* is a preposition, *þæt* a subjunction. *Œþ* is AUX to *þæt*. It also exists as one word (which had to be split in the tokenization, because B&T does not list it as a single entry, although CH does).

• *œþ þe*: 91102

• *sona swa*: 118730. *swa* is head, *sona* is AUX

• *swa swa*: ‘as’ 118687.

• *swa þæt*: ‘so that’ 118740. *þæt* is head, *swa* is AUX.

• *to þam þæt*: 118677, 91209 ‘in order that’. *to* and *þam* are AUX to *þæt*. Of course, this type of analysis suppresses the fact that *þam* gets dative case from the preposition *to*. An alternative would be to analyze *þam* as OBL to *to* and *þæt* as APOS to *þam*. This analysis would, however, suppress the fact that the expression is grammaticalized.

• *to þon þæt*: 116694.

• *þa þa*: ‘then when’ (often followed by a main clause introduced by *þa*): The first *þa* is AUX to the second *þa*. The first *þa* is an adverb in the morphology, the second is a subjunction.

• *þa hwile þe*: 116664, 116688, 91162

• *þæs þe*: 118850, 107852, 90896, 91149

• *þeah þe*: 107814. *þeah* is AUX to *þe*. *þeah* is an adverb in the morphology, *þe* is the subjunction.

• *þe læs*: 102495

• *þe læs þe*: 119022

• *þi læs þe*: 119082

3.14 **COMPLEX CONJUNCTION**

There is one complex conjunction as well, namely *hweþer þe*, as found in 102426 and 102370. The tree structure in (57) shows the rather complex analysis of 102426. (The XSUB slash from *forspillan* to *sawla* is unfortunate – but it is not possible to get an OBJ slash there.)

(57) *alyfð restedagum wel to donne hweþer ðe yfele, sawla gehælan, hweþer ðe forspillan*

(102426)

permits restdays-DAT well to do whether ðe evil, souls heal, whether ðe destroy ‘is it lawful to do good on the sabbath days, or to do evil? to save life or to kill?’
3.15 **BOTH ... AND**
In constructions with *both...and, both* is AUX to *and*. See Menotec 9.2.2.

3.16 **NEGATION**
The negation *ne* is analyzed as AUX in all cases. In cases where it is cliticized onto the verb (*næs ne + wæs, nat ne + wat*), the token is split in the analysis, to e.g. *n + at*. When *naht/noht/neaht* (from *nawiht*) is used as an adverb meaning ‘not’, it is tagged as an adverb (non-inflecting) in the morphology, and gets the syntactic function AUX, like *ne*. The lemma form is *naht*.

3.17 **IMPERSONAL CONSTRUCTIONS**
In impersonal constructions such as (58), the experiencer, here *pe*, is analyzed as OBL. Impersonal verbs are not ‘visible’ in the syntax, unless there is an expletive pronoun.

(58) *ac gif ðu wille witan hwæt he sy, axsa hine, forðam þe gedafenað þæt þu wite*  
(118926)  
but if you will know what he is, ask him, because you behave that you know-SUBJ.  
‘but if you wish to know what he is, ask him, for it is fitting that you should know’
See 118771 for another example of an impersonal construction.

3.18 EXISTENTIAL SENTENCES AND EXISTENTIAL THERE

There is reason to believe that existential sentences, of the type *There is a man in the street*, existed already in Old English, because OE has constructions that are exactly like PdE, i.e. sentences with *there* + a locative adverbial (Breivik, 1981). However, it was also common to repeat elements, so in cases where the clause contains *there* and a locative adverbial, *there* might also be an adverbial. The question therefore arose as to how to analyze such sentences. The simplest solution would be to analyze *there* and the locative adverbial as two adverbials, with the locative adverbial as an APOS to *there*. However, existential sentences are interesting from a word order and information structure point of view, because it was in this type of sentence V2 remained the longest – until late ME (and relics still exist in PdE: *On the table was a book*). It is therefore a good idea to make them retrievable in the corpus as far as possible. The analysis is therefore as follows: If a sentence contains *there* + V + NP + a locative adverbial, *there* is tagged morphologically as an adverb (since the adverb is the origin of existential *there*). In the syntax, the NP is SUB and *there* is tagged as EXPL(itive), and a dependent to PRED. If the sentence has *there* + V + NP and no locative adverbial, *there* is tagged as ADV and dependent on PRED. Thus, the diagnostic for existential *there* is that there must also be a locative adverbial present in the sentence. A researcher interested in constructions with existential *there* may search for EXPL constructions (but will then also get
expletive *hit*, see section 3.19). A researcher interested in constructions with *there* in general may search for the lemma *þær*. Note that only existential sentences with existential *there* are easily retrievable in this way; existential sentences without *there* are harder to find, but a query including the verb *be* and an adverbial may be a start.

(59) & *þær* sint swiðe micle meras fersce geond *þa* moras (115628)
    and there are very large lakes fresh beyond the moors
    ‘and there are very large fresh water lakes beyond the moors’

Sentence 116683 is another example of existential *there*, and example (31) above.

3.19 EXPLETIVE HIT
In sentences such as *hit* gelamp *þæt*... ‘it happened that...’, *hit* is analyzed an an EXPL(etive) element, and the *þæt* clause is a COMP. *Hit* is also an expletive in *þonne hit dæg was* ‘when it was day’. In the syntax, *dæg* is XOBJ, and the slash goes back to the verb, since *hit* is non-referential. In the morphology, *hit* is a neuter pronoun in the nominative case.

3.20 PERSONAL PRONOUN AS REFLEXIVE
In OE, the simple personal pronouns served as reflexive pronouns (though constructions with *self* were beginning to be used, cf. section 3.21). These pronouns are analyzed as OBL (if dative) or OBJ (if accusative) in the syntax. (60) is an example.

(60) *þa* bewende he hine & cidda Petre & cwað (102884)
    then turned he him(self) and rebuked Peter and said
    ‘then he turned around and rebuked Peter and said’
3.21 SELF

Self is tagged as an adjective in the morphology (cf. YCOE guidelines). In constructions such as Crist sylf ‘Christ self’ (107917), sylf is used for emphasis (cf. Quirk & Wrenn, 1989: 72), and annotated syntactically as an APOS to Crist. In constructions such as heo hyre silf ‘she herself’ (118942), silf is APOS to heo and hyre is ATR to silf.

There are some OE constructions that resemble the PdE reflexive. An example is 118844: Mid þi þe he þas þingc was sprecede to him silfum ‘while he was saying these things to him self’. These are analyzed in the same way as when self is used for emphasis. Him is OBL to the head of the PP to, and silfum is APOS to him. Likewise, in 18304, (he) ætywde hine sylfne ‘(he) showed himself’, hine is OBJ to ætywde, and sylfne is APOS to hine. In short, self is always tagged as APOS, whether it is used for emphasis or in a manner that resembles the modern reflexive.

3.22 THAT IS (id est)

There was some discussion in our group (since this is a construction that also occurs in the Romance languages) as to how to analyze the appositive conjunct (cf. Quirk et al., 1985: 635) that is (Latin id est). At first we analyzed the that is part as PARPRED – a parenthetical predicate (see section 3.34). For example, in a sentence like 107822 (þa) wæron eunuchi, þæt synt belisnode ‘they were eunuchs, that is castrated’ (the OE verb and participle are in the plural), þæt synt would be PARPRED under the root, and castrated would be APOS under eunuchs. However, we did not find this analysis very satisfactory, because it seemed strange to separate the þæt synt part from the complement. We therefore ended up analyzing the entire ‘that is’ construction as an APOS; so in 107822, þæt synt belisnode is APOS to eunuchi. Further examples of this type of construction can be found in 115601, 102754, 102805, 102751, 91080.

---

6 There is a problem with this analysis, too, namely the fact that belisnode does not agree with þæt, so it is slightly strange to have belisnode as XOBJ to þæt. Normally, the XOBJ element gets case from the element to which it is a complement.
However, there is one type of that is construction that has to be analyzed as a PARPRED, namely the type exemplified in 107827-107828: [Uia iustorum recta facta est, et iter sanctorum preparata est:] pet is, para rihtwisra weg is gerihtleced ‘[Latin text:] that is, “the way of the righteous is guided’’. Here pet is must be attached to the root as PARPRED.

### 3.23 PARTITIVE GENITIVE

We use the relation PART for partitive genitives. An example is *fela fægera þinga* ‘many wonderful things’ (118968) in which *fægera þinga* are in the genitive case (*fela* takes genitive). *Fela* is the head, *þinga* is PART, and *fægera* is ATR to *þinga*. Another example is *ælc para þinga* ‘each of the things’ (118969): *ælc* is head, *þinga* is PART and *para* is ATR. In 118748, *hwæt is manna þe nyte þæt...* ‘what man is there who knows not that...’ *manna* is a genitive plural, and PART to *hwæt*. The Menotec project decided to make do without PART, for the sake of simplicity, and annotate such elements as ATR. Note that in the ‘modern’ expression’ with an of-genitive, we use ATR, and not PART, cf. e.g. 102922.

### 3.24 DIRECT AND INDIRECT QUESTIONS

In direct and indirect questions, the question word is an argument of the verb. It may be an XOBJ, as in 102470: *hwylc is min modor & mine gebroþru?* ‘who is my mother and my brethren’ and 118919: *hwæt is þes iunga man* ‘who is this young man’ (see also example (35) above), or an OBL, as in 118842: *Hwæs mæg ic biddan?* ‘for what can I beg?’, and 118999: *Leofe dohtor, for hwi eart ðu þus ærwa col?* ‘Dear daughter, why are you thus early awake?’.

### 3.25 SWA HWÆT SWA, SWA HWILC SWA

In constructions such as *swa hwæt swa* (e.g. 118972), *swa hwilc swa* (e.g. 102655, 107857), *swa hwilc NP swa* (e.g. 102654, 118678), the two *swas* are analyzed as AUX elements, basically for want of a better way of doing it. In 118678, for example: *swa hwilc man swa minne rædels riht aræde* ‘whatever man reads my riddle right’, the second *swa* might possibly be a sort of relative pronoun: ‘whatever man who reads my riddle right’, which would yield a different analysis. However, for the sake of simplicity, these *swa* constructions have all been analyzed in the same way. YCOE seems to take the second *swa* as a complementizer.

### 3.26 ARG

Normally, accusative objects are tagged as OBJ and oblique objects as OBL. However, if the form of the NP does not make it clear whether the case is accusative or something else (genitive or dative), the ARG tag is used, even if we know from the form of the sentence what type of element the NP most probably is. An example is found in (61), where we cannot tell what case *gretinge* is in, since feminine nouns have an -e ending both in the accusative and the dative case, and there is nothing else in the NP that reveals the case of the head. In such cases, the tag ARG(ument) is used in the syntax. There are numerous examples of this in the corpus.

(61) *Da forseah he Apollonius cyrlisces mannes gretinge æfter ricra manna gewunan* (118766)
then despised he Apollonius humble-GEN man-GEN greeting after rich-GEN men-GEN custom

‘Then he, Apollonius, despised the greeting of a humble man, after the custom of rich men’

3.27 NARG

The tag NARG is used for nominal arguments, typically in cases where there is a deverbal noun, of the type fear of, comfort of, forgiveness of, praise of... (See PROIEL guidelines, section 9.5 and Menotec guidelines section 7.3.6.) Examples from the corpus are *Iohannes was on westene fulligende & bodiende dædbote fulwiht on synna forgynenesse* ‘John was in the wilderness, preaching the baptism of repentance for the remission of sins’ (102277), where the deverbal noun is *fulwiht* ‘baptism’, and the NARG element is *on synna forgynenesse*; and 107856: *...mid goddre gedrohtnunge to Godes þeowdome* ‘with good devotion to God’s service’, where the deverbal noun is *gedrohtnunge* ‘devotion’, and the NARG element is *to Godes þeowdome*.

Note that NARG is only used with nouns. If an adjective takes an argument, it is tagged as OBL, cf. e.g. 119021, 118817. See also 107922, where the complement to the adjective is a finite subclause (‘worthy that...’), and thus a COMP.

3.28 TWO OBL ARGUMENTS

When there is a dative and a genitive argument, the sentence gets two OBL arguments, as in (62).

(62) *and geond twæntig daga hyre næs getyðod æniges bigleofan binnan ðam þeostrum* (107984)
and throughout twenty days her-DAT not-was granted any-GEN nourishment-GEN in the darkness ‘and for the space of twenty days, she was given no food in the darkness’

3.29 **INFinitive of INTension/PURPOSE = XADV**
Infinitives of intention/purpose are tagged as XADV, as in (63). *Secan* is XADV to *gan*.

(63)  *and gan we secan ure gesthus þæt we magon us gerestan* (118987)
and go we seek our guesthouse that we may us rest ‘and let us go seek our guesthouse, so that we may rest’
3.30  FOCUS ADVERB

In (64), *eac* is a focus adverb to *biscop*, and gets the function ADV in the syntax. In other words, here we get an ADV dependent to a noun (*biscop*). The focus adverb functions pragmatically rather than to restrict the reference of the NP. Another example is found in 118757: *na þæt an his find, eac swilce his frind* ‘not only his enemies, but also his friends’.

(64)  *Pa wurdon gelome pa leofan mædenu Eugenia and Basilla, and eac se biscop, on sunderspræce swiðe gebys gode* (107950)

then were frequently the dear maidens Eugenia and Basilla, and also the bishop, in private-speech much employed

‘Then the dear maidens and also the bishop were frequently employed in private conversation’

3.31  ABSOLUTE CONSTRUCTIONS

Old English texts that are translated from Latin, especially Bible texts, often try to emulate the Latin absolute constructions. Absolute constructions are embedded predications with an adverbial relation to the rest of the sentence (see PROIEL guidelines, section 10.7), i.e., the participle is the head, and is related to the matrix clause via the relation ADV. The dative element is the SUB(ject) of the participle verb. In (65), the absolute construction is *gefultumigendum Gode*. The participle *gefultumigendum* is an ADV to the matrix verb *eom*, and *Gode* is the SUB to the participle.

(65)  *ac for eowre gesælde gefultumigendum Gode ic eom hider cumen* (118819)

but for your advantage help.PRS.PTCP.DAT God.DAT I am hither come

‘but for your advantage, with the help of God, I have come here’
In (66), the absolute construction is *hyra handa gegripenre*, with *gegripenre* as a dependent on the matrix verb, and *hyre handa* as SUB to the participle.

(66) & *genealæcende he hi up ahof hyre handa gegripenre* (102329)
    and approaching he her up lifted her hand-DAT taking-PRF.PTCP.DAT
    ‘and approaching, he lifted her up, taking her hand’
The corresponding Latin example is (67). Note that the Latin participle is passive.

(67) *et accedens elevavit eam adprehensa manu eius* (PROIEL 10342)
and approaching lifted her taking-PRF.PTCP.ABL hand-ABL her
‘and approaching, he lifted her up, taking her hand’

![Diagram of sentence structure]

The last example is (68), where the first *him* is the SUB of the participle *gangendum* in the absolute construction *him of scipe gangendum*. The corresponding Latin example can be found in PROIEL 10552.

(68) & *him of scipe gangendum* *him sona agen arn an man of þam byrgenum on unclænum gaste* (102554)
and him-DAT from ship walk-PRS.PTCP.DAT him immediately towards ran a man from the tombs with unclean spirit
‘and when he had come out of the ship, there ran towards him from the tombs a man with an unclean spirit’
Further examples of absolute constructions can be found in PROIEL 102344, 102613, 102649, 102722, 107925, 118859, 102640, 118795, 102887, 103031, 103365, 103399, 103495, 103614, 103638, 103657. As regards gewordenum restedæge, it corresponds to PROIEL 10626, but note that PROIEL 10626 is analyzed as an impersonal absolute construction. They take the meaning to be ‘it became sabbath’, not ‘sabbath became’, so that sabbato is an XOBJ in an impersonal construction. For OE, however, I chose to analyze it as a regular absolute construction, since it seems that the OE scribes interpreted it as such with the rendering gewordenum.DAT restedæge.DAT, which is the typical rendering of absolute constructions. The reading is thus ‘when the sabbath came’.

3.32 SOURCE AND GOAL ARGUMENTS
Source and goal arguments are analyzed as OBL, in clauses of the type He came from Galilea and went to Jerusalem.

3.33 BUTAN
Butan can be either a subjunction or a preposition. If it is a preposition, the prepositional complement is analyzed as OBL, as in ac he ne mæg for scame in gan buton scrude ‘for shame, he may not go in without clothing’ (118901). The PP may have different functions, but it is usually ADV. If butan is a subjunction, it introduces an adverbial subclause (ADV), as in Pharisei & ealle Iudeas ne etað buton hi hyra handa gelomlice þwean ‘The Pharisees and all the Jews do not eat unless they wash their hands’ (102752). Note that there may be cases where butan looks like a preposition, but the case of the complement tells us that it cannot be, and the construction therefore has to be rendered with an empty verb. An example
is (69), in which *God ana* is in the nominative case. It can therefore not be the complement of the preposition, and is instead analyzed as the subject in a verbless adverbial clause.

(69)  *Hwa mæg synna forgyfan buton God ana?* (102367)

who may sins forgife except God alone

‘Who may forgive sins except God alone?’

3.34  **PARPRED**

Parenthetical predicates are of the following types:

1) inserted comment clauses in direct speech, as in 116666 *Nat ic, cwæd Orosius, hwæder mare wundor wæs...* ‘I do not know, said Orosius, which was the greater wonder...’ *Cwæd Orosius* is a PARPRED here. In 102371, the PARPRED, *He cwæd to þam laman* ‘He said to the lame’ is surrounded by direct speech;

2) impersonal constructions such as *him þuhte* ‘him (dative) thought’ in 115591;

3) *þæt is* constructions of the type in 107828 (see section 3.22);

4) inserted direct speech, as in 102699: *and þæt þu cwæde mooddrenum flesce ic bruce, ne eart þu on þam leogende* ‘and what you say “on maternal flesh I feed”, in that you are not
lying’. Here, on _moddrenum flesce ic bruce_ is a PARPRED. See also 102370 for several inserted direct speech sequences analyzed as PARPRED.

### 3.35 Topicalization
It is quite common in Old Germanic languages to use topicalization to establish what the sentence is about. An example is in the story of Ohthere, when Ohthere tells the king that he went northwards primarily for the walrus tusks, which were very valuable. Then the story goes on:

(70) þa teð hie brohton sume þæm cyninge (115593)  
the teeth they brought some the-DAT king-DAT  
‘The brought the king some of the teeth’

As we see, the syntactic analysis is not able to render the topicalized structure here – _þa teð_ is an APOS to _sume_. However, it is possible to find this kind of structure in queries if one searches for APOS elements that occur _before_ their head, since other APOS constructions will always occur _after_ their head. Although word order is not visible in the trees, every token has an index number, which means that word order is retrievable. Another interesting case of topicalization is found in 116691.

### 3.36 _FORDAM_ as conjunction
_Fordam_ is usually a subjunction (‘because’), or an adverb (‘therefore’), but it can also be a conjunction meaning ‘for’. An example is (71):

(71) _ga onbæc Satanas forþam þu nast þa ðing þe synd Godes ac þa ðing þe synd manna_ (102885-102886)  
go behind Satan for you not-know the things that are God-GEN but the things that are men-GEN  
‘get behind, Satan, for you do not know the things that are of God, but the things that are of men’
In the syntactic analysis, *forðam* is analyzed as AUX, in the same way as other conjunctions.

Other examples are 102781 and 122410.

### 3.37 SOPLICE and WITODLICE

We distinguish between *soplice* and *witodlice* used as sentence adverbials or regular adverbials. If they are sentence adverbials, they are regarded as being outside the clause structure, and in our application they thus get the tag VOC. VOC stands for ‘vocative’, which is clearly not an ideal tag, but it was the only possible option in the application. If *soplice* or *witodlice* are not sentence adverbials, they get the tag ADV. (72) is an example of *soplice* as a sentence adverbial, and (73) is an example of *soplice* as an adverbial.

(72) *soðlice he is dead* (102951)
    verily he is dead
    ‘verily, he is dead’
(73)  And Dryhten soðlice heofonas geworhte (107820)
and Lord truly heavens created
‘And the Lord truly created the heavens’
4 REFERENCES


Haugland, Kari. 2007. Old English Impersonal Constructions and the Use and Non-use of Nonreferential Pronouns. Dr.art. dissertation, University of Bergen.


