



A CORPUS-BASED APPROACH TO THE LEMMATISATION OF OLD ENGLISH SUPERLATIVE ADVERBS¹

YOSRA HAMDOUN BGHIEL
Universidad de La Rioja
yohamdou@unirioja.es

ABSTRACT. *The aim of this article is to discuss the lemmatisation process of Old English adverbs inflected for the superlative from a corpus-based perspective. This study has been conducted on the basis of a semi-automatic methodology through which the inflectional forms have been automatically extracted from The York-Toronto-Helsinki Parsed Corpus of Old English Prose and The York Toronto-Helsinki Parsed Corpus of Old English Poetry whereas the task of assigning a lemma has been completed manually. The list of adverbial lemmas amounts to 1,755 and has been provided by the lexical database of Old English Nerthus. Additionally, the resulting lemmatised list has been checked against the lemmatised forms compiled by the Dictionary of Old English and Seelig's (1930) work on Old English comparative and superlative adjectives and adverbs. Through this comparison it has been possible to verify doubtful forms and incorporate new ones that are unattested by the YCOE. This pilot study has implemented for the first time a methodology for the lemmatisation of a non-verbal class and can be further applied to those categories that are still unlemmatised, namely nouns and adjectives.*

Keywords: Old English, lemmatisation, superlative adverbs, corpus linguistics, lexicography.

¹ This research has been founded by Grant PID2020-119200GB-I00 funded by MCIN/AEI/10.13039/501100011033.

UN ENFOQUE BASADO EN CORPUS PARA LA LEMATIZACIÓN DE LOS ADVERBIOS SUPERLATIVOS DEL INGLÉS ANTIGUO

RESUMEN. Este artículo tiene como objetivo analizar el proceso de lematización de los adverbios en grado superlativo del inglés antiguo desde un enfoque basado en corpus. Este estudio se ha realizado siguiendo una metodología semiautomática mediante la cual el proceso de extracción de formas flexivas de *The York-Toronto-Helsinki Parsed Corpus of Old English Prose* y de *The York Toronto-Helsinki Parsed Corpus of Old English Poetry* ha sido automático, mientras que la tarea de asignación de lema ha sido llevada a cabo manualmente. El listado de lemas consta de 1.755 formas y ha sido proporcionado por la base de datos léxica del inglés antiguo *Nerthus*. Asimismo, los resultados obtenidos del proceso de lematización han sido cotejados con las formas atestiguadas por el *Dictionary of Old English* y por la obra de Seelig (1930) sobre los adjetivos y adverbios en grado comparativo y superlativo del inglés antiguo. A través de este estudio comparativo ha sido posible verificar formas dudosas e incorporar nuevas atestigüaciones que no han sido registradas por el YCOE. Este análisis piloto ha implementado por primera vez una metodología para la lematización de una clase no verbal que puede aplicarse al resto de categorías que todavía no han sido lematizadas, en concreto los nombres y los adjetivos.

Palabras clave: inglés antiguo, lematización, adverbios superlativos, lingüística de corpus, lexicografía.

Received 20 February 2020

Revised version accepted 27 July 2020

1. AIMS AND RELEVANCE OF THE RESEARCH

This article deals with the lemmatisation of the Old English superlative adverbs. Lemmatisation is understood as the process by which a group of words are morphologically related and reduced to a lemma or headword, including both the predictable and the unpredictable forms. Lemmatisation is a necessary task in the field of Old English lexicography, and of historical linguistics in general, as there is not, for the moment, a fully lemmatised corpus or dictionary that compiles a systematic inventory of the Old English lexicon. Several reasons explain this fact; one of the most convincing arguments that explains the inconsistent spelling variation is the diversity of dialects (Kentish, West Saxon, Mercian and Northumbrian) and the absence of a written standard of the language.

One of the pending tasks historical corpus linguistics must tackle is precisely to increase the amount of grammatically annotated material. However, as Claridge (2008: 254) points out, this proves quite challenging in a historical language as Old

English due to aspects such as form and usage differences and the high internal variability. In addition to this, the limited amount of Old English texts that have been preserved, together with the difficulties in accessing some of this material, which may exist exclusively in the form of a manuscript, reduce the possibilities of finding evidence of less frequent grammatical phenomena.

This study follows one of the current lines of research of the *Nerthus* Project, which centres on the lemmatisation of the Old English lexicon. A few studies have already been conducted that tackle the lemmatisation of Old English verbs, including strong verbs (Metola Rodríguez 2015, 2017), preterit-present, anomalous and contracted verbs (García Fernández 2018, fc.) and weak verbs (Tío Sáenz 2019). Although substantially different methodologies, in all three cases the lemmatising task involves a semi-automatic procedure that requires an eventual manual revision and evaluation of the results.

Within this context, the pilot study presented here is conceived as a first attempt to lemmatise a non-verbal category, more specifically adverbs inflected for the superlative. A methodology has been devised to this aim that can be applied to the rest of categories that are not fully lemmatised yet, these are nouns and adjectives.

This work mainly contributes to identifying a lemma for all the inflected forms that are tagged as superlative adverbs by the *York-Helsinki Parsed Corpus of Old English Prose* (Taylor *et al.*, eds., 2003) and *Poetry* (Pintzuk and Plug, eds., 2001). Furthermore, the comparison of the results with other sources has enriched the study by providing mutual feedback. One of these sources is *The Dictionary of Old English* (DOE), only available for letters A-I, which represents a valuable lexicographical source as it attests a vast amount of inflectional forms and has been elaborated on the basis of a corpus containing one example, at least, of every surviving text in this language. The other is a secondary source, the work by Seelig (1930). It consists of a compilation of adjectival and adverbial forms in the comparative and the superlative that have been grouped together under the adequate lemma.

This study is grounded on the fields of lexicography and corpus linguistics, two disciplines that bear an intimate relationship due to the mutual feedback they obtain. It is common knowledge among the scholars of this area that a quality lexicographical work must be necessarily based on a corpus. Many scholars have addressed this close connection. Faaß (2017), for instance, attaches utmost importance to the contribution of corpus linguistics to electronic dictionaries as it permits the analysis of language in natural contexts. Besides, this author draws attention to the fact that corpora should be digitalized for the sake of a greater

accuracy in data description (Faaß 2017: 124). Last but not least, another key aspect is the convenience of having fully annotated corpora, namely morphologically, syntactically and semantically, that render a valuable tool for lexicographical purposes.

Given this background, the remainder of the paper is organized as follows. Section 2 summarizes the main information regarding the description of the adverbial system and their formation. Section 3 offers a description of the sources that have been utilised in both the extraction and lemmatisation processes. Section 4 deals with the methodological steps that have been followed and the decisions made in the lemmatisation process. Section 5 discusses the results of the analysis. The paper closes with the main conclusions of the study.

2. AN OVERVIEW OF THE OLD ENGLISH ADVERBIAL SYSTEM

Old English is the term coined to allude to the vernacular language spoken in Great Britain and attested in manuscripts previous to 1100. It is a Germanic language, more specifically a West Germanic one, sharing group with other languages such as Old High German, Old Saxon and Frisian.

The period known as Old English comprises around 600 years (5th-11th centuries), although the majority of texts that are preserved date back to the late 9th and 11th centuries. Four main dialects existed at that time, namely Kentish, West Saxon, Mercian and Northumbrian. It is precisely the presence of this variety of dialects that makes Old English spelling system particularly complex. Besides the coexistence of different dialects, the abundant vernacular literary activity also contributed a significant variation in orthography and inflexion. In this context, West-Saxon was instituted as the standard written language as most of the manuscripts of that time belong to this dialect.

This language is characterized by a rich inflectional system with a close relation holding between word form and word function (Smith 2009: 22). Mitchell and Robinson, in turn, prefer to describe Old English as a “half-inflected language” (1985: 62) as it preserves only four cases of the eight that were present in Indoeuropean. Moreover, prepositions are often found introducing phrases in an oblique case, a circumstance in which the use of the preposition seems unnecessary.

As a synthetic language, the function of a word is marked through inflectional endings. In the case of adverbs, the only inflection they may undergo is that corresponding to the comparative and the superlative degrees. The typical

adverbial endings expressing degree are *-or* and *-ost*, although there are also other possibilities, these are *-ar* and *-ur* for the comparative and *-ast*, *-est* and *-ust* for the superlative.

The adverbial lexicon represents roughly a five percent of the total of Old English words. Adverbs amount to 1,755, a considerable low figure if compared with nouns, for example, which constitute half of the Old English lexicon, or with verbs and adjectives, adding up to twenty percent each. As in Present Day English, Old English adverbs were used as modifiers of adjectives, verbs and other adverbs.

Adverbs were mostly created through the addition of suffix *-e* to an adjective, which corresponds to the ending that indicated the instrumental case. For instance, *glæd* 'joyous, glad' > *glæde* 'joyously, gladly'; *sār* 'sore' > *sāre* 'sorely'. Ending *-e* was lost by the end of the 14th century, as a result a number of adjectives and adverbs are formally alike. An example is *ānwille*, which is both an adjective ('obstinate, stubborn') and an adverb ('wilfully, obstinately'); something similar occurs with *dēore*, meaning 'dear, beloved' as an adjective, and 'dearly; with kindness' when it is an adverb.

Many adverbs originate in adjectives ending in *-lic*, which have been added suffix *-e* eventually. For instance, adjective *nāhtlic* 'worthless' gave rise to adverb *nāhtlice* 'worthlessly'; in like manner, *rȳnelic* 'mystical' creates adverb *(ge)rȳnelice* 'mystically', hence *-lice* became a typical ending that was appended to adjectives to form adverbs (*rōtlice* 'glad, cheerfully' < *rōt* 'glad, cheerful').

A few adverbs, and also adjectives, form the superlative through a double suffixation process (Fulk 2018: 240). In these cases, there is an *-m-* right before the superlative suffix. For instance, both *eallmāest* 'almost' and *endemest* 'equally' undergo double suffixation. In this regard, Campbell (1959: 278) observes that this double suffixation is especially common in inflected adjectives that originate in an adverb.

A number of adverbs experience mutation in the root vowel when they form the comparative or superlative, being this mutation the only evidence of degree. An example is *eald* 'old' – *yldra* (comp.) – *yldest* (superl.). A few adverbs undergo suppletive comparison, that is, the stems of the positive and of the comparative or superlative forms is not the same; this is the case, for instance, of *wel* 'well' – *bet* (comp.) – *betst* (superl.). Attending to their formation process, adverbs can be divided into two groups: basic and non-basic. Regarding the first group, no productive morphological process is involved. Examples of basic adverbs, also called 'monomorphemic' (Fulk 2018: 237), are *nŪ* 'now', *oft* 'oft', *under* 'under', *yme* 'around', etc. Non-basic adverbs constitute roughly 90% of the

adverbial lexicon. Non-basic adverbs include zero derived/converted², affixed and compound adverbs³. Zero derivation affects, approximately, a 10% of the adverbs. This process consists of category extension and semantic modification without formal change. Examples of zero derived adverbs are *tōbeald* (adv.) ‘inclined, forward, in advance manner’ < *tōbeald* (adj.) and *unnēab* (adv.) ‘not near, far, away from’ < *unnēab* (adj.). The most numerous group of adverbs, around 70% of the total, encompasses affixed adverbs. These have been created either through prefixation or suffixation. Examples of affixed adverbs are *sāmbāl* ‘unwell, weakly’ (< *bēal* ‘well, safe, whole’), *ymbūtan* ‘around, about, outside, beyond’ (< *ūtan* ‘from outside’), *sorglice* ‘miserably, grievously’ (< *sorglic* ‘sorrowful, miserable’) and *wynnnum* ‘joyfully, beautifully’ (< *wynn* ‘pleasure, gladness’). Finally, compound adverbs amount to sixty-four and have been created by joining two lexeme stems of the same or different category. Compound adverbs are, for instance, *ādunweard* ‘downwards’ (< *ādūn* ‘down’ + *weard* ‘towards’) and *hysewīse* ‘like young men’ (< *hyse* ‘young man’ + *wīse* ‘wise, manner’).

A group of adverbs are subject to suppletive comparison (Fulk 2018: 240), which means that the resulting comparative and superlative forms have a different stem from that of the positive adverb. Examples of adverbs undergoing suppletive comparison are *yfle* ‘evil’ > *wiers* - *wierst*; *wel* ‘well’ > *bet/sēl* - *bet(e)st/best/sēlest*.

3. SOURCES

The study presented here is framed within the *Nerthus* Project, one of whose current research interests is the lemmatisation of the Old English lexicon. To do so, different sources are required.

To start with, the York-Toronto-Helsinki Parsed Corpus or York Corpus of Old English (henceforth YCOE), which is divided into prose and poetry, has provided the data of analysis. The YCOE is a 1.5 million-word corpus which is both morphologically and syntactically annotated. The annotation draws on the system used by the Penn-Helsinki Parsed Corpus of Middle English. Figures 1 and 2 below showcase the fragment annotated *Saga me hwilc sunu wræce ærest his fæder on hys moder innoðe* [Ad:12.1.31_ID] both morphologically and syntactically:

2 The difference between zero derivation and conversion lies in the fact that the former entails a formal contrast between the source and the derived term that has an effect on the whole inflectional paradigm, while in the case of conversion the rest of the paradigm is not necessarily affected (Martín Arista 2011).

3 All the information related to the classification and quantification of adverbs per formation process draws on Maíz Villalta (2012).

```
<T06110002400,12.1>_CODE $Saga_VBI me_PRO hwilc_WADJ^N sunu_N^N
wr+ace_VBDS +arest_ADVS^T $his_PRO$ f+ader_N on_P hys_PRO$ moder_N^G
inno+de_N^D ._. coadrian,Ad:12.1.31_ID
```

Figure 1. POS (Part of Speech) annotation in the YCOE.

```
(IP-MAT-SPE (VBI $Saga)
  (NP (PRO me))
  (CP-QUE-SPE (WNP-NOM-1 (WADJ^N hwilc) (N^N sunu))
    (IP-SUB-SPE (NP-NOM *T*-1)
      (VBDS wr+ace)
      (ADVP-TMP (ADVS^T +arest))
      (NP (PRO$ $his) (N f+ader))
      (PP (P on)
        (NP-DAT (NP-GEN (PRO$ hys) (N^G moder)
          (N^D inno+de)))))) (.) (ID
      coadrian,Ad:12.1.31))
```

Figure 2. PSD (Parsed) annotation in the YCOE.

As aforementioned, the YCOE constitutes the source of the data. The inflectional forms (superlative adverbs in this case) together with their tags have been extracted from this corpus in order to be lemmatised. At this juncture it must be remarked that the totality of the inflected forms belongs to *The York-Helsinki Parsed Corpus of Old English*, as no evidence of superlative adverbs has been found in its poetry counterpart.

Besides, in order to lemmatise forms, it has been necessary to have a reliable list of lemmas “under which elements of the corpora containing the word forms of the same lexeme are represented” (Burkhanov 1998: 122). This list has been retrieved from *Nerthus*, a lexical database of Old English storing around 30,000 predicates free of context marking. This database draws on three main sources: Hall-Meritt’s *A Concise Dictionary of Anglo-Saxon* (and the *Supplement*) constitutes the primary basis as it has guided the spelling of both the headwords and the alternative spellings. To a lesser extent, *Nerthus* is based on *An Anglo-Saxon Dictionary* by Bosworth-Toller⁴ and on Sweet’s *The Student Dictionary of Anglo-Saxon*.

4 Including Toller’s *Supplement* and Campbell’s *Addenda*.

Three fields of *Nerthus* database are of interest for this study, these are predicate, spelling variants and predicate translation. The field “spelling variants” does not contain independent predicates or morphologically contrastive words but variants of the predicate they refer to (Martín Arista 2010: 10). The field “predicate translation” provides an equivalent of the source term in Present Day English as offered by the standard dictionaries of reference. Figure 3 below illustrates how this information is offered by the *Nerthus* database:

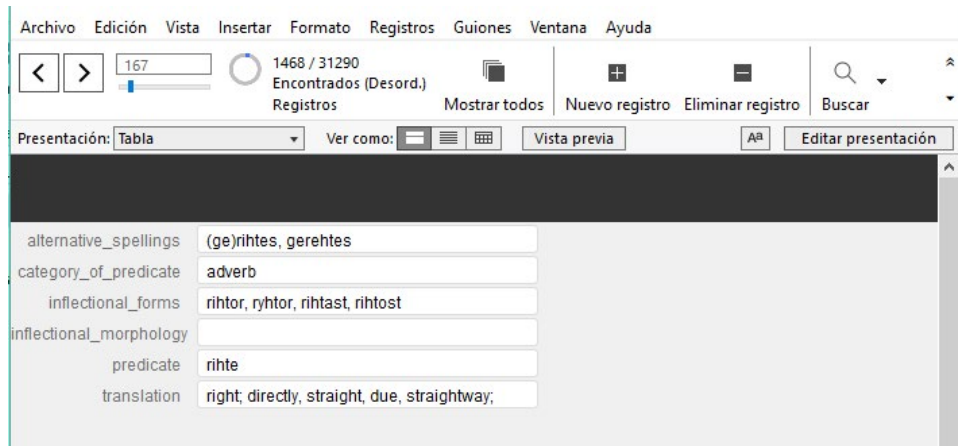


Figure 3. Fields extracted from *Nerthus* database.

The lemmatisation task has required the support of other sources, especially in those cases in which lemma assignment was particularly complicated. Seelig's (1930) work, titled *Die Komparation der Adjektiva und Adverbien im Altenglischen*, has contributed to both verifying and completing the analysis. This work basically consists of a compilation of Old English adjectives and adverbs in the comparative and the superlative, as attested in a number of texts listed at the beginning of the book, and organized under a common lemma. The second chapter of this work addresses adverbs and is divided into three subchapters on a morphological basis. Firstly, Seelig brings together all the adverbs that undergo regular comparison, being the most numerous group. The second group gathers all the adverbs whose comparative entails root vowel change. An example is adverb *ēaðe*, *ēðe*, *iðe* 'easy', which has attested the comparatives *æð*, *ēað*, *ēð*, *iēð*, *ið* and *yð* and the superlatives *ēaðost*, *ēaðust*, *ēaðusð*, *ēðest*, *iðesð*, *yðæst*, *yðast* and *yðost*. The third group of adverbs is composed of those subjected to irregular comparison, i.e. the comparative and superlative are created out of a different stem from that of the positive. Examples of irregular comparison are adverbs *wyrse* 'worse' (< *yfle* 'bad') and *bet* 'better' (< *wel* 'well').

The Dictionary of Old English (henceforth DOE) has been the main lexicographical source of reference in this study as it is, for the time being, the most complete source in what regards the number of attested inflectional forms in each entry as it is based on a collection of texts that comprise, at least, one copy of every surviving text in Old English. However, there is a major drawback, which is the fact that the DOE has only published headwords starting with A-I. The entries in the DOE include grammatical information of the headword, namely part of speech, gender and grammatical class. In addition to the different attested spellings of the headword in the corpus, each entry also provides the inflectional forms, dialectal variations, the number of occurrences in the corpus and the meaning accompanied by textual citations.

The Dictionary of Old English Corpus (henceforth DOEC) is an essential reference for scholars interested in the study of Old English as it compiles over 3 million words of this language in a more than 3,000 texts belonging to the years 600-1150. As stated in the previous paragraph, it is the source that supplies the dictionary with written evidence about the lexical stock of the Old English language. A simple search in the corpus allows to obtain all the hits corresponding to the searched word, together with the short title of the text in which it appears and the Cameron number.

4. METHODOLOGY

Lemmatization is commonly understood as “the reduction of inflectional word forms to their lemmata, i.e. basic forms, and the elimination of homography” (Burkhanov 1998: 122). According to this author, this process “involves the assignment of a uniform heading under which elements of the corpora containing the word forms of same lexeme are represented” (1998: 122).

The lemmatization process is divided into three main stages: extraction, lemma assignment and validation. This process is highly dependent upon the approach adopted: in a full-form approach, the decision of lemma assignment falls on the author, whereas from a paradigm-based perspective, words that are paradigmatically related share the same lexeme or lemma. This research adopts the second methodology, as the objective is to gather both the predictable and unpredictable forms under the same lemma.

To start with, the first step in the lemmatization process is the extraction of the material that will be lemmatized next. This material consists of a list of adverbial lemmas and a list of inflectional forms. The former list has been automatically retrieved from the Old English database *Nerthus* by exporting the field called “predicate” exclusively of those predicates that are adverbs. Additionally, two other

fields have been exported, these are “spelling variants” and “predicate translation”, as they contain helpful information to carry out this task. By way of illustration, Table 1 includes a few adverbial predicates together with their alternative spellings, if any, and the translation.

Table 1. List of adverbial lemmas extracted from *Nerthus*.

Predicate	Alternative Spellings	Predicate Translation
<i>cēne 2</i>	<i>cýne 2</i>	<i>in warlike wise</i>
<i>(ge)bēotlīce</i>		<i>in a threatening manner, threatingly (BT)</i>
<i>(ge)blīðe 2</i>		<i>joyfully, gladly</i>
<i>(ge)brægdēnlīce</i>	<i>(ge)bregdenlīce</i>	<i>cunningly, deceitfully (DOE)</i>
<i>(ge)cwēmīce</i>		<i>graciously, kindly, humbly, satisfactorily</i>
<i>(ge)cyndelīce</i>		<i>naturally (DOE)</i>
<i>(ge)dīegollīce</i>	<i>(ge)dēgollīce, (ge)dēagollīce, (ge)dēogollīce, (ge)dýgollīce, (ge)dīgollīce (BT), (ge)dīgolīce (BT), (ge)dēgelīce, (ge)dēglīce, (ge)dēgullīce, (ge)dīegellīce, (ge)dīgelīce, (ge)dīgle, (ge)dīhlīce</i>	<i>secretly; softly (of the voice)</i>
<i>(ge)ðwæreīce</i>	<i>(ge)ðwærīce (BT)</i>	<i>in accord (Sweet)</i>
<i>(ge)dwoīce</i>		<i>foolishly, heretically; erroneously, ignorantly, stupidly (BT)</i>
<i>(ge)ðyldelīce</i>	<i>(ge)ðyldigīce (BT), (ge)ðyldelīce (BT)</i>	<i>patiently, quietly</i>
<i>(ge)ðyldigīce</i>		<i>patiently (Sweet)</i>
<i>(ge)dyrstigīce</i>	<i>(ge)dyrstelīce</i>	<i>boldly, daringly (BT)</i>

Regarding the list of inflectional forms, this has been retrieved from the YCOE. This corpus is annotated with both POS (Part-of-Speech) and PAS (parsed) labels. At this stage of the process, only the morphological information, i.e. POS labels, will be used. To be more precise, the list of inflectional forms consists of all the forms which have been tagged with the label ADVS (superlative adverb).

The methodology for the systematic extraction of all the attested inflectional forms corresponding with the label ADVS begins by launching a preliminary search on the POS files. Next, the POS file is opened in the Notepad ++ text editor to follow a step by step procedure to ensure uniformity during the extraction procedure. Then, a number of adjustments are made: symbols +a, +d and +t are replaced with æ, ð and þ, respectively; spaces are replaced with mark paragraphs,

resulting in a column which is arranged in alphabetical order; punctuation marks, text codes and other undesired information are eliminated. Afterwards, the list that has been obtained divided in texts is copied into a single Excel file. Inflectional form and tag are separated into two columns and two additional fields are appended to the right, these are the text code and the text genre⁵ (prose or poetry). By storing the material in this way, it is possible to have access to a comprehensively organized inventory of forms that allows for both quantitative and qualitative analyses. Searches can be also launched that are restricted by tag, text code or genre, permitting a variety of approaches in this regard. A sample is provided in Table 2.

Table 2. Sample of extracted superlative adverbs from the YCOE.

Inflectional form	Tag	Text	Genre
<i>æryst</i>	ADVS^T	cowsgosp.o3	PROSE
<i>andgitfullicast</i>	ADVS	coboeth.o.02	PROSE
<i>andgitfullicost</i>	ADVS	Coprefcura	PROSE
<i>beorbtost</i>	ADVS	coboeth.o.02	PROSE
<i>firrest</i>	ADVS^L	coorosiu.o2	PROSE
<i>fyrrest</i>	ADVS	cocuraC	PROSE
<i>fyrrest</i>	ADVS^T	codocu3.o3	PROSE
<i>nybst</i>	ADVS^T	comart3.o23	PROSE
<i>nybst</i>	ADVS^L	coorosiu.o2	PROSE

As observed in the second column, there are some tags that further specify the type of adverbial: ADVS^T stands for temporal superlative adverb and ADVS^L means locative superlative adverb. In the table above, the forms *æryst* and *fyrrest* are examples of the first type, whereas *firrest* is a locative. This information proves particularly useful when two forms are alike but have different tags. Consider the case of *fyrrest*, a form that may be just a superlative or convey a temporal meaning. Something similar occurs to the form *nybst*, which may be locative ('near, nigh; about, almost, nearly') and temporal ('lately').

Once the material is conveniently arranged, the next stage of the process consists of assigning a lemma to the inflectional forms. This process is yet far from being automatic. Table 3 below illustrates how it takes place. As a novelty, an additional column has been inserted to the left of the table that includes the corresponding lemmas. In the event that the lemma list does not provide any adequate lemma for a particular form, the corresponding cell is left blank.

5 It must be noted that no examples of superlative adverbs have been found in poetry texts.

Table 3. A sample of lemmatised superlative adverbs.

Lemma	Inflectional Forms	Tag	Text	Genre
<i>beorbte</i>	<i>beortost</i>	ADVS	coboeth.o.02	PROSE
<i>bet</i>	<i>best</i>	ADVS	colaw2cn.o3	PROSE
<i>beorbte</i>	<i>biorbtost</i>	ADVS	coverhom	PROSE
<i>beorbte</i>	<i>biorbtust</i>	ADVS	coverhom	PROSE
<i>fæstlice</i>	<i>fæsðlicost</i>	ADVS	cocuraC	PROSE
<i>fæste</i>	<i>fæstost</i>	ADVS	coaelive	PROSE
<i>feorr</i>	<i>firrest</i>	ADVS^L	coorosiu.o2	PROSE
<i>fullice</i>	<i>fullecost</i>	ADVS	cocuraC	PROSE
<i>fullice</i>	<i>fullicost</i>	ADVS	cocuraC	PROSE

The list of inflectional forms is arranged alphabetically. In this way, all the tokens that are formally alike appear together, facilitating the task of assigning a lemma. The lemmatisation of the Old English superlative adverbs displaying regularity does not present much difficulty as the possible spelling variations are, for the most part, attested in the “alternative spellings” column from the lemma list. In a first round roughly eighty percent of the inflectional forms were assigned a lemma whereas twenty percent remained unlemmatised. These unlemmatised forms evinced some type of irregularity, therefore it was necessary to resort to additional sources to facilitate this task. The sources consulted include *The Dictionary of Old English*, available for letters A-I, which contains a fairly complete list of attested forms in each entry; also Bosworth and Toller’s dictionary, although in this case the amount of inflectional forms attested is considerably inferior to the former dictionary. Besides, the main Old English grammars of reference, especially Campbell (1959), have been highly clarifying in this regard.

The third and final task in the lemmatisation process involves a contrastive analysis with the aforementioned sources. This step of the process has two main objectives: on the one hand, it permits the verification of the associations established between lemmas and inflectional forms assigned to them; on the other hand, it allows for the completion of the process in those cases in which no lemma has been assigned to a particular form or when new inflectional forms are suggested.

To illustrate this final stage of the process, Table 4 represents two different adverbial lemmas, *æf* ‘previously, before that, soon, formerly’ and *ēadelīce* ‘easily’, and the inflectional forms (adverbs in the superlative) assigned to them by the YCOE, the DOE and Seelig.

Table 4. Attested inflectional forms belonging to the lemma *ær* in the different sources.

Lemma	YCOE	DOE	Seelig (1930)
<i>ær</i>	<i>ærcæst, ærcast, æresð, ærest, ærost, ærst, ærust, æryst, æst, arest, erest, erost</i>	<i>ærest, æresð, ærest, æreast, æryst, ærcæst, ærcast, ærost, ærust, æst, ærst, arest, aryst, erest, erost, eræst, earest. hærest, æres, ærets, ærist, ærist, æris, ærst, aræst,</i>	<i>ærcast, ærcæst, æres, ærest, æresð, æris, ærist, ærst, ærest, ærust, æryst, æst</i>
<i>ēaðelīce</i>	<i>eaðelicost eaðelicust, eðelicost</i>	<i>eaþelicost, eaðelicost; eaðelicust, eaðelucust, eþelicost, eðelicost, æðelicest</i>	<i>eaðelicost, eaðelicust eðelicost</i>

As observed in this table, the DOE is the source that attests a considerably higher amount of inflectional forms. Starting with lemma *ær*, the DOE provides ten new forms that are not attested by the YCOE, these are *eræst*, *earrest*, *hærest*, *æres*, *ærets*, *ærist*, *aerist*, *æris*, *ærst* and *aræst*. To a lesser extent, Seelig also contributes with new forms, in this case these are *æres*, *æris* and *ærst*. Regarding lemma *ēaðelīce*, something similar occurs. Quantitatively speaking, the DOE is the source that gathers the most complete list of inflectional forms, a total of seven. In this case, all the forms attested by the YCOE are also verified by the DOE. The DOE, in addition, provides evidence of four new forms, these are *eaþelicost*, *eaðelucust*, *eþelicost* and *æðelicest*. Two of these forms contain the letter *þ* (*eaþelicost* and *eþelicost*), creating pairs of forms such as *eaðelicost* and *eaþelicost* or *eðelicost* and *eþelicost*. This distinction is, however, not present in either the YCOE nor Seelig. As for Seelig, this author does not attest any new form for lemma *ēaðelīce*.

5. ANALYSIS AND RESULTS

In this section, the methodology previously described is applied to the superlative adverbs retrieved from the YCOE. In the first place, an overview of the results of the extraction process will be presented. Next, the lemmatised forms will be compared with the ones provided by the DOE (only for A-I adverbs) and by Seelig. The last part of this section will deal with doubtful forms and the decisions made in this regard.

A total of 1,270 forms⁶ have been tagged as superlative adverbs according with the POS files furnished by the YCOE. Once these forms have been identified, the list of lemmas of the *Nerthus* database will allow for the completion of the lemmatising task. This list contains 1,755 adverbial lemmas, although only 80 have turned out to be lemmas for these forms.

The task of lemma assignment has been accomplished manually. Each inflected form is addressed individually and the most suitable lemma from the *Nerthus* list is allocated. Once adverbs are deprived of the corresponding inflectional endings, these are *-æst*, *-ast*, *-esð*, *-est*, *-ost*, *-st*, *-ust*, *-yst*, *-exð*, *-osð*, *-(e)-m-est*, *-or*, *-estan*, *-usð*, the resulting form does not differ much from the adverbial lemma. In fact, this procedure has succeeded in assigning a lemma to nearly all the inflected forms, yet the following seven remained unlemmatised in the first round: *edost*, *leofost*, *liffest*, *liofast*, *suiðusð*, *suiðust*, and *ytemest*. These forms present a higher degree of opaqueness, resulting in doubtful or ambiguous cases that required deeper research. In a few cases, it was simply found that the morphological analysis of the YCOE was not correct. Although the list of lemmas provided by *Nerthus* has mostly proved to be suitable, it was necessary to consult other sources when it was unclear which lemma to assign. In order to fill this information gap, *The Dictionary of Old English* and Seelig's (1930) work turned out to be of valuable help.

An initial analysis of the data evinces that most of the lemmas are associated with just one inflectional form, being the number of lemmas assigned to several forms much more limited. In the case of superlative adverbs, a total of forty lemmas hold a one-to-one correspondence with an inflectional form. Examples of hapax legomena include (the lemma is given in brackets): *onlicost* (*onlice* 'like, in like manner, similarly'), *fægerost* (*fægre* 'softly, pleasantly, gently'), *orenlicost* (*orenlice* 'excessively, immoderately') and *fæsðlicost* (*fæstlice* 'firmly, certainly, stoutly'). A considerably smaller quantity of instances has been found for hapax dislegomena, that is, a lemma that is assigned to two inflectional forms. Examples of this type are *rihtost*, *ryhtosð* (*rihte* 'right; directly, straight'), *eapelicost*, *eðelicost* (*eaðelice* 'easily'), *niedemesð*, *nyðemyst* (*niðer* 'down, donwards, beneath, below') and *andgitfullicast*, *andgitfullicost* (*andgietfullice* 'intelligibly, sensibly, clearly, plainly, distinctly'). Diametrically opposed to these cases are lemmas *swiðe* and *ær*. Lemma *swiðe* 'very, much, exceedingly; severely; fiercely, violently; chiefly, especially, mostly' has been assigned to fifteen different forms (types) that add up to 221 tokens. Twelve

⁶ The Appendix contains the full inventory of lemmas that have been used in the analysis together with the inflectional forms lemmatised under each headword. Information about their textual occurrence is also included.

inflectional forms share lemma *ær* ‘previously, before that; soon; formerly, before’, which amount to 690 tokens. Both lemmas also coincide in the rich spelling variation of their inflectional forms. The following forms have been gathered under lemma *swiðe*: *swiðast*, *swiðest*, *swiðosð*, *swiðost*, *swiðost*, *swiðusð*, *swiðust*, *swiðast*, *swiðest*, *swyðost*, *swyðust*, *swyðpast*, *swyðor*, and *swyðost*. Of these forms, it is *swiðost* the one with the greatest amount of occurrences in the YCOE, a total of ninety-seven. As for *ær*, the forms gathered under this lemma include *arest*, *ærast*, *æræst*, *ærerst*, *ærersð*, *ærrost*, *ærst*, *ærust*, *æryst*, *æst*, *erest*, *erost*. Among these forms, *ærerst* turns out to be the form with the highest number of occurrences, a total of 601.

Some inflected adverbs receive a specific tag that further specifies the type of comparative or superlative. Regarding the superlative, apart from ADVS, which indicates superlative degree, there are two other tags, the locative ADVS^L and the temporal ADVS^T. The YCOE has allocated the tag ADVS to 422 adverbs, ADVS^L to twenty-five forms and ADVS^T to 762. As evinced by the figures, superlatives with a temporal meaning substantially outnumber the rest of categories.

On the other hand, the set of inflectional adverbs provided by the YCOE, although constituting a fairly comprehensive list, does not collect the totality of Old English superlative adverbs. This is the reason why this study has been enhanced with the comparison of forms attested by each of these three sources, namely the YCOE, the DOE and Seelig. Table 5 below represents the procedure adopted in the lemmatisation of superlative adverbs and the subsequent comparison of sources.

Table 5. Comparing the results of the lemmatisation with other sources.

Lemma	Inflectional form	Tag	Text code	Text genre	Seelig		DOE	
					Lemma	Inflectional form	Lemma	Inflectional form
<i>andgietfullice</i>	<i>andgietfullicost</i>	ADVS	<i>Coprefcura</i>	PROSE	✓	X	X	✓
<i>beorhte</i>	<i>biorhtost</i>	ADVS	<i>Coverbom</i>	PROSE	✓	X	✓	✓
<i>beorhte</i>	<i>beorhtost</i>	ADVS	<i>Conicoda</i>	PROSE	✓	✓	✓	✓
<i>fæstlice</i>	<i>fæsðlicost</i>	ADVS	<i>cocuraC</i>	PROSE	✓	✓	✓	✓
<i>bāte</i>	<i>batust</i>	ADVS	<i>Colacnu.o23</i>	PROSE	✓	✓	✓	X

As observed in the table above, two columns have been added that correspond with information relating to the DOE’s and Seelig’s attestations. Symbol ✓ has been used to indicate that both the lemma and the inflectional form are attested by a source and have the same status, i.e. of lemma or of inflectional form of that lemma. Symbol X indicates that either the lemma or the inflectional form are not attested by any of the sources; this same symbol may also reveal that a

different lemma has been assigned to an inflectional form. By way of illustration, lemma *andgietfullīce* has been attested by Seelig, although this author does not list the inflectional form *andgitfullicost*. The opposite is true for the DOE; this dictionary attests the form *andgitfullicost*, however the lemma that is assigned in this case is *andgytfullīce*, an alternative spelling of the *Nerthus*' lemma *andgietfullīce*. *Beorbte* is a lemma in both sources, the only difference is that one of its inflectional forms, *biorbtost*, is not attested by Seelig. In the fourth line, it can be observed that both the lemma and the inflectional forms are included with the same status in Seelig and the DOE. Finally, *bāte* is a lemma in both sources, however *batust* is unattested by the DOE; it appears, however, in the DOEC as a unique occurrence.

The analysis has confirmed that there is a total of 886 inflectional forms (tokens) starting with letters A to I that have been mapped into eighty lemmas. Of the total, 858 tokens (70 types) have been attested by the DOE and 768 tokens (36 types) by Seelig. Some inconsistencies found in the lemmatisation of superlatives have to do with the assignment of dissimilar lemmas by the different sources. An example of this is the form *fægerost*, which has been assigned *Nerthus*' lemma *fægre* 'fairly, elegantly', although the DOE opts for *fægere*. In other cases, the DOE has helped identify a lemma for those forms undergoing suppletive comparison, such as *fyrrest* (*feor* 'far, far away').

Another possible problem is the existence of two potential lemmas for one form. This occurs with *innest*. In the headword list of reference there are two lemmas, namely *inn* 'in, into, inwards, within, inside of; inwardly' and *inne* 'in, inside, within, in-door', that can be associated to the inflectional form. A search for the form in the DOEC revealed that there is only one occurrence in the corpus, particularly in the text from the *Cura Pastoralis*; the citation in question is *tihð his fet sua he **innest** mæg* (CP B9.1.3 [1149 (35.241.7)]), which is used by the DOE under the entry for *in*, *inn*. Consequently, the form *innest* could be lemmatised under lemma *inn* on the basis of the DOE's criterion.

This study has discovered two inflectional forms, *andgitfullicost* and *eðost*, which have not been attested by the DOE. The lemma suggested by *Nerthus* are *andgietfullīce* 'sensibly, clearly, plainly, distinctly; intelligibly (BT)' and *ēaðe* 'easily, lightly'. In these particular cases, although the form is unattested by the DOE, this dictionary attests a lemma that has the status of alternative spelling in *Nerthus*. This analysis has also permitted the identification of forms with the status of lemma in the DOE but which are part of the inflectional paradigm in *Nerthus*. This is the case of *bet*, which is an example of suppletive comparison of the adverb *wel* 'well, abundantly; very, very easily, very much' in this database.

The study has also identified a divergence in what regards the morphological status of the following forms *eallmæst*, *endemest* and *fyrst*. The YCOE has tagged them as inflectional adverbs in the superlative and they have been assigned *Nerthus*' lemmas *eallmæst* 'nearly all, almost, for the most part', *endemest* 'equally, likewise, at the same time, together' and *fyrst* 'at first', respectively. These forms are considered attested spellings rather than inflectional forms of their corresponding headwords in the DOE.

As for the inflectional forms starting with letters L to W, they add up to 293 tokens (36 types), which have been mapped into forty lemmas. By comparing the results with the forms compiled by Seelig, it can be checked that the vast majority of them have been attested by the author. There is one form that has a different status in *Nerthus* and in Seelig's work, this is *wyrs*. According to Seelig it is a lemma, whereas in *Nerthus* this form is part of the inflectional paradigm of *yfel* 'evil, ill, wicked, bad, wretched'.

This research has contributed seventeen superlative forms that have not been compiled by Seelig, although their lemma is attested by the author. These forms are: *geornlicast*, *geornlicest* (*geornlice* 'openly, manifestly'), *gewissost* (*wise* 'wisely'), *batust* (*bate* 'hotly'), *healicast* (*bēallice* 'highly, aloft'), *bluddost* (*blūde* 'loudly, aloud'), *længast*, *længest* (*lange* 'long'), *nearwlicast* (*nearolice* 'narrowly, closely'), *raðust* (*braðe* 'hastily, quickly'), *ribtlicost* (*ribtlice* 'justly, uprightly'), *swiðest*, *swiðosð* (*swiðe* 'very much'), *teartlicost* (*teartlice* 'sharply, severely'), *ðwyrlicost* (*ðwēorlice* 'insolently'), *ungeredelicost* (*ungerædllice* 'sharply, roughly'), *widdast* (*wide* 'widely').

The manual assignment of a lemma is not always a straightforward task. It might be the case that there is no formal coincidence in the lemmas suggested by the sources consulted. In this context, the lemma suggested by *Nerthus* prevails over the rest. To illustrate this, the form *fægerost* has been associated to *Nerthus*' lemma *fægre*, a form that is attested as an alternative spelling of *fægere* by the DOE.

As occurred with *inmest*, a similar circumstance has been identified in the assignment of a lemma to the forms *suiðusð* and *suiðust*. Apparently, there are two possible options in the lemma list, these are *suið* 'southwards, south' and *sið* 'late, afterwards'; in this case, however, there is a clear difference of meaning in both candidates. The solution adopted here was to verify the meaning of *suiðusð* and *suiðust* in the citations where they appear: *ðonne ðonne hie hie selfe **suiðusð** eaðmedað* (CP B9.1.3 [1457 (41.301.14)]) 'Then they humbled themselves the latest' and *ðeah ða tunga **suiðust** mænde* (CP B9.1.3 [1517 (43.309.8)]) 'Still the tongues declare the latest'. As observed, the suitable lemma is *sið* 'late, afterwards' in both cases.

Three other inflectional forms required closer inspection as their lexical category is not adverbial but adjectival, these are *leofost*, *liffest* and *liofast*. The fact that neither *Nerthus* nor Seelig offered a suitable lemma arouse suspicion, that is why it was deemed appropriate to verify their lexical status in context. The following citations correspond to the occurrences of these forms in the DOEC: *þonne hit wære leofost gebealden* (WHom 13 B2.3.1 [0004 (12)]) ‘and often it is more quickly lost when it is held dearest’; *min bearn liffest gedoan* (Ch 1510 (Rob 6) B15.6.27 [0002 (4)]) ‘my child has done the quickest’; *swæ him liofast sie* (Ch 1510 (Rob 6) B15.6.27 [0004 (11)]) ‘as it may best please them’. As evinced in the examples, they perform an adverbial function, however these are the only three occurrences that have been tagged by the YCOE as superlative adverbs, the rest being adjectives. This leads us to suggest that even if the function they fulfil is adverbial, they are adjectives. This hypothesis is supported by Bosworth and Toller’s dictionary, which confirms their adjectival status.

Likewise, *endenexð*, which has been tagged as a superlative adverb by the YCOE, is an adjective. This conclusion was reached after verifying its status in different sources. Firstly, *Nerthus* does not provide any lemma that could be assigned to this form. In addition, the DOE considers this form an alternative spelling of the adjectival headword *endēnext*.

Overall this section has presented the quantitative and qualitative results of the lemmatisation of the Old English superlative adverbs as attested by the YCOE. The lemmatising task has been fully accomplished for the totality of forms and new ones have been identified that have completed the original inventory. Finally, a comparative analysis with the DOE and Seelig has provided mutual feedback and has helped verify and refine the results of the analysis.

6. CONCLUSIONS

This article has contributed to the design and implementation of a methodology for the lemmatisation of Old English adverbs inflected in the superlative. If compared with the verbal lexicon, already lemmatised in previous works by the *Nerthus* project, adverbs present a substantially lower degree of variation and opacity, which has motivated a different methodology, more appropriate for this class.

The lemmatising methodology has consisted of three main tasks: firstly, the extraction of the 1,267 forms from the YCOE identified with the tag ADVS, which has been fully automatic; secondly, the lemmatisation of these forms through the manual assignment of a lemma from the list of headwords supplied by *Nerthus*;

finally, the comparison of the results with a lexicographical and a secondary source. This three-step procedure has thus evinced that the study is based on a combination of sources and of analytical methods that are yet far from being completely automatic.

This work has also laid bare the difficulties that this process entails and has suggested solutions. Basically, the difficulties encountered are in line with the presence of opaque forms, ambiguous forms that can be assigned more than one lemma and, to a lesser extent, forms that have been originally wrongly analysed and are not adverbs. In order to tackle these issues, additional sources have been consulted, such as Old English grammars and dictionaries of reference in the language; in other cases, doubtful forms have been analysed in context to determine their meaning and function as attested in citations.

Considering the work that has been previously developed by members of the *Nerthus* group and the study presented here, it is possible to make further advances in the lemmatisation of the pending major categories, including nouns, adjectives and non-graded adverbs. In addition, a work of these characteristics has direct implications in the field of corpus linguistics as it has proved the feasibility of lemmatising a historical corpus.

7. REFERENCES

- Bosworth, J. and T. N. Toller. 1973 (1898). *An Anglo-Saxon Dictionary*. Oxford: Oxford University Press.
- Burkhanov, I. 1998. *Lexicography. A Dictionary of basic terminology*. Wydawn: Wyższej Szkoły Pedagogicznej w Rzeszowie.122.
- Campbell, A. 1959. *Old English Grammar*. Oxford: Oxford University Press.
- Claridge, C. 2008. "Historical corpora". *Corpus Linguistics. An international Handbook*. Eds. A. Lüdeling, M. Kytö and T. McEnery. Berlin: Mouton de Gruyter. 242-254.
- Faab, G. 2017. "Lexicography and corpus linguistics". *The Routledge Handbook of Lexicography*. Ed. P. A. Fuertes-Olivera. Abingdon: Routledge. Accessed 15 May 2019, Routledge Handbooks Online. 123-425.
- Fulk, R. D. 2018. *A Comparative Grammar of the Early Germanic Languages*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 237-240.
- García Fernández, L. 2018. "Preterite-present verb lemmas from a corpus of Old English". *Verbs, Clauses and Constructions: Functional and Typological Approaches*. Eds. P. Guerrero Medina, R. Torre Alonso and R. Veá Escarza. Newcastle: Cambridge Scholars Publishing. 59-76.

- García Fernández, L. (fc.). *The lemmatisation of the verbal lexicon of Old English on a relational database. Preterite-present, contracted, anomalous and strong VII verbs.* PhD Dissertation, Department of Modern Languages, University of La Rioja.
- Hall, J. R. C. 1984 (1960). *A Concise Anglo-Saxon Dictionary.* With a supplement by H. D. Meritt. Toronto: University of Toronto Press.
- Healey, A. diPaolo *et al.* (ed.). 2018. *Dictionary of Old English: A to I online.* Toronto: Dictionary of Old English Project, Centre for Medieval Studies, University of Toronto.
- Healey, A. diPaolo with John Price Wilkin and Xin Xiang. 2009. *Dictionary of Old English Web Corpus.* Toronto: Dictionary of Old English Project, Centre for Medieval Studies, University of Toronto.
- Maíz Villalta, G. 2010. *The Formation of Adverbs in Old English: A Functional View.* Master's Dissertation, Department of Modern Languages, University of La Rioja.
- Martín Arista, J. 2010. "Building a Lexical Database of Old English: issues and landmarks". *Current Projects in Historical Lexicography.* Ed. John Considine. Newcastle: Cambridge Scholars Publishing. 1-33.
- Martín Arista, J., L. García Fernández, M. Lacalle Palacios, A. E. Ojanguren López and E. Ruiz Narbona. 2016. *Nerthus V3. Online Lexical Database of Old English.* Nerthus Project. Universidad de La Rioja [www.Nerthusproject.com]
- Metola Rodríguez, D. 2015. *Lemmatisation of Old English Strong Verbs on a Lexical Database.* PhD Dissertation, Department of Modern Languages, University of La Rioja.
- Metola Rodríguez, D. 2017. "Strong Verb Lemmas from a Corpus of Old English. Advances and issues". *Revista de Lingüística y Lenguas Aplicadas* 12: 65-76.
- Mitchell, B. and F. Robinson. 1985 (1964). *A Guide to Old English.* Oxford: Blackwell.
- Pintzuk, S. and L. Plug (eds.). 2001. *The York-Helsinki Parsed Corpus of Old English Poetry.* Department of Language and Linguistic Science, University of York.
- Seelig, F. 1930. *Die Komparation der Adjektiva und Adverbien im Altenglischen.* Heidelberg: Winter.
- Smith, J. J. 2009. *Old English. A Linguistic Introduction.* Cambridge: Cambridge University Press.
- Sweet, H. 1976 (1896). *The Student's Dictionary of Anglo-Saxon.* Cambridge: Cambridge University Press.
- Taylor, A., A. Warner, S. Pintzuk and F. Beths (eds.) 2003. *The York-Toronto-Helsinki Parsed Corpus of Old English Prose.* Department of Language and Linguistic Science, University of York.
- Tío Sáenz, M. 2019. "A semi-automatic lemmatisation procedure: Old English class 1 and class 2 weak verbs". *Electronic Lexicography of Old English* 1: 67-76.

8. APPENDIX. LIST OF LEMMAS AND LEMMATISED FORMS

Ēr ‘previously, before!’: *ærst* [Cocathom1.O3, Coeluc2], *æræst* [Coboeth.o.02], *ærast* [Coboeth.o.02, CochronD], *æresð* [Coboeth.o.02, CochronD], *ærest* [Coadrian.O34, Coaelhom.O3, Coaelive.O3, Coalcuin, Coalex.O23, Coapollo.O3, Cobede.O2, Cobenrul.O3, Coboeth.o.02, Cobyrhtf.O3, Cocanedgx, Cocathom1.O3, Cocathom2.O3, Cochdrul, Cochronc, CochronD, Cocura.C, Codocu1.04, Codocu3.O23, Codocu3.O3, Coeluc1, Coexodusp, Cogregdc.O24, Cogregdh.O23, Coherbar, Coinspold.O34, Coinspolx, Colacnu.O23, Colaece.O2, Colaw1cn.O3, Colaw2cn.O3, Colaw5atr.O3, Colaw6atr.O3, Colsigewz.O34, Colwgeat, Comart2, Comart3.O23, Comary, Coneot, Conicoda, Conicode, Coorosiu.o2, Coinspold.O34, Coinspolx, Colsigewz.O34, Conicoda, Coorosiu.o2, Covinsal, Cowulf.O34], *ærost* [Cogregdc.O24, Conicodc].

Andgietfullice ‘sensibly, clearly, plainly, distinctly, and intelligibly’: *andgitfullicast* [Coboeth.o.02], *andgitfullicost* [Coprefcura.O2].

Beorbte ‘brightly, brilliantly, splendidly; clearly, lucidly, distinctly’: *beorbtoſt* [Coboeth.o.02, Conicoda, Coverhom], *beortoſt* [Coboeth.o.02], *biorbtoſt* [Coverhom], *biorbtuſt* [Coverhom].

Beorblice ‘splendidly’: *gebeorblicost* [Coverhom].

Bet ‘better, of manner better’: *beſt* [Colaw2cn.O3], *betesð* [CocuraC], *betest* [CocuraC], *betſt* [Coaelholm.o3, Coaelive.O3, Coapollo.O3, Coboeth.O2, Cobyrhtf.O3, Cochdrul, CochronD, CocuraC, Codicts.O34, Codocu3.O3, Coinspolx, Colaw2cn.O3, Colaw5atr.O3, Colaw6atr.O3, Conicodc, Coorosiu.o2, Coverhom, Cowulf.O34].

Clæne ‘clean, clearly, fully, purely’ Superlative: *clænoſt* [Cowulf.O34].

Ēaðe ‘easily, lightly, soon; willingly, readily’: *eaðoſt* [Cowulf.O34, Colaece.O2], *eaðuſt* [Cowulf.O34], *eðeſt* [Coalcuin], *iðeſð* [CocuraC], *yðeſt* [coeuphr].

Ēaðelice ‘easily (BT)’: *eaðelicost* [coherbar], *eaðelicuſt* [Cowsgosp.O3], *yðelicor* [Coalex.o23, Cogregdc.o24], *eðelicost* [Coboeth.o.02].

Eallmæſt ‘nearly all, almost, for the most part’: *eallmæſt* [Cochronc, CochronD].

Endemest ‘equally, likewise, at the same time, together, unanimously; fully, entirely; in procession’: *ændemest* [Coboeth.o.02], *endemest* [Coboeth.o.02].

Ende-næxt ‘last, final’: *endenexð* [Coaelholm.o3].

Fægge ‘fairly, elegantly, beautifully; pleasantly, softly, gently, kindly; well, justly; early’: *fægeroſt* [Coverhom].

Fæſte ‘fast, firmly, securely; straitly, strictly; heavily (sleep); speedily’: *fæſtoſt* [Coaelive.O3].

Fæſtlice ‘firmly, constantly, fast, quickly’: *fæſðlicost* [CocuraC].

Feorr ‘far, far away, distant, remote; far back (in time); further, besides, moreover’: *feorst* [Cogregdc.O24], *firrest* [Coorosiu.o2], *fyrrest* [Coboeth.o.02].

Forbæfendlice ‘continently’: *forbæfendlicust* [Cochdrul].

Fulllice ‘entirely, fully, perfectly, completely’: *fullecost* [CocuraC], *fullicost* [CocuraC].

Fyrmest ‘at first, most, very well, best’: *fyrmest* [Coaelive.O3, Coboeth.o.02, Cocathom, Cochronc, Cochronc, CocuraC, Codocu3.O3, Cogenesic, Coherbar, Coinspolx, Colaeece.O2, Colaw2cn.O3, Colaw6atr.O3, Colwgeat, Coorosiu.o2, Cootest.O3, Coverhom, Cowulf.O34], *fyrmust* [Coherbar].

Fyrst ‘at first’: *fyrst* [Cochad .024].

Gearwe ‘well, effectually, sufficiently, thoroughly, entirely; quickly; near’: *gearost* [Cogregdc.O24].

Gebende ‘near, at home; closely, in detail’: *gebendost* [Coaelive.O3, Cocathom1, Contempo.O3].

Gelice ‘like as if’: *geliccost* [Coverhom], *gelicost* [Coalquin, Coboeth.o.02, CocuraC, Codicts, Coleofri.O4, Cogregdc.O24, Colaeece.O2, Coorosiu.o2, Coverhom], *gelicust* [Cowsgosp.O3].

Gelōmlīce ‘often, repeatedly (Sweet)’: *gelomlicost* [Coorosiu.o2].

Georne ‘eagerly, zealously, earnestly, gladly; well, carefully, completely, exactly; quickly’: *geornast* [Coinspolx], *geornest* [Coalquin], *geornost* [CocuraC, Coinspold.O34, Coinspolx, Colaw1cn.O3, Colaw2cn.O3, Colaw5atr.O3, Colaw6atr.O3, Cowulf.O34].

Geornlice ‘earnestly, diligently, zealously, strenuously, carefully, willingly’: *geornlicast* [Coinspolx], *geornlicest* [Cowulf.O34].

Gesundiglice ‘prosperously (BT)’: *gesundlicost* [Coboeth.o.02].

Grimlice ‘fiercely, severely, cruelly (BT)’: *grimlicost* [Comart3.O23].

Hūte ‘hotly, fervidly’: *batost* [Colacnu.O23, Colaeece.O2], *battost* [Coverhom], *batust* [Colacnu.o23].

Hēalīce ‘highly, aloft; in or to high position or rank, loftily; intensely, very’: *healicast* [Cobenrul.O3].

Hearde ‘hard, hardly, firmly, very severely, strictly, vehemently; exceedingly, greatly; painfully, grievously’: *heardost* [Coboeth.o.02, Coorosiu.o2].

Hlūde ‘loudly, aloud’: *bluddost* [Coverhom].

Hraðe ‘hastily, quickly, promptly, readily, immediately, soon; too soon’: *braðost* [Coboeth.o.02, Cocathom1, Coherbar], *raðer* [Cowulf.O34], *raðosð* [CocuraC],

raðost [Cochronc, Cochronc, Cogenesic, Coinspolx, Colaw2cn.O3, Colwstan1.O3, Coorosiu.o2, Cootest.O3, Cosevensl, Cowulf.O34], *raðust* [Cowsgosp.O3].

Inn ‘in, into, inwards, within, inside of; inwardly’: *inmest* [CocuraC].

Lāðe ‘inimically, in detestation’: *laðost* [Colaw6atr.O3].

Lange ‘long, a long time, far’: *længast* [Codocu3.O3], *længest* [Cogregdc.O24], *længst* [Cochronc].

Late ‘late; slowly; at last; lately’: *latost* [Cobenrul.O3, Coboeth.o.02, Cogregdc.O24].

Leng ‘longer’: *lengest* [Cochronc, Codicts, Coorosiu.o2], *lengst* [Colsigewz.O34].

Līðe ‘in a gentle, soft, calm, mild; gracious, kind, agreeable, sweet manner’: *licost* [Coorosiu.o2].

Mærlīce ‘gloriously, splendidly; excellently’: *mærlīcost* [coaelive.o3, cowulf.O34].

Mæst ‘most, chiefly, especially’: *mæst* [Cochronc, Cochronc], *meast* [Cochronc].

Nēah ‘near, nigh; about, almost, nearly, lately’: *neaxst* [Coboeth.o.02], *neahst* [Coboeth.o.02, Comart3.O23], *nebst* [Cochronc, Cogregdc.O24, Colaw1cn.O3], *nebeste* [Coboeth.o.02], *nest* [Coelofri.04], *nexð* [Cochronc], *next* [Coboeth.o.02], *nibst* [Coorosiu.o2], *nybst* [Comart3.O23, Coorosiu.o2].

Nearolīce ‘narrowly, closely, briefly, accurately; strictly, stringently, oppressively; evilly’: *nearwlicast* [Coinspolx].

Nīðer ‘below, beneath, down, downwards’: *niedemesð* [CocuraC], *nyðemyst* [Contempo.o3].

Nytwierðlice ‘in a ‘useful, profitable, manner’: *nyttwyrdlicost* [CocuraC], *nyttweorðlicor* [CocuraC].

Oft ‘above, on high; to or on the other side; from side to side, across; beyond, above (quantity)’: *oftosð* [Coboeth.o.02, CocuraC], *oftost* [Coaelive, Coboeth.o.02, Cocathom2, Cochronc, CocuraC, Coinspold.O34, Coinspolx, Colaeece.O2, Colaw6atr.O3, Colwsigexa.O34, Coorosiu.o2, Coprefcura, Coquadru.O23, Coverhom, Covinsal, Cowulf.O34], *oftust* [Coherbar, Coverhom].

Onlice ‘in a like, resembling, similar manner’: *onlicost* [Coorosiu.o2].

Orenlice ‘openly, manifestly, plainly, clearly, unreservedly’: *orenlicost* [Coverhom].

Ribte ‘right, due, straight (of direction, as in right on, due east), outright; precisely, exactly, just; rightly, duly, well, correctly, truly, properly, fairly, justly; directly, immediately’: *ribtost* [Coinspold], *ryhtosð* [CocuraC].

Ribtlice ‘justly, uprightly, virtuously; properly, rightly, regularly; correctly, precisely’: *ribtlicost* [Cobyrtf, Cocathom1, Codocu3.O3, Coinspold.O34, Coinspolx, Contempo.O3].

Scortlice ‘shortly, briefly, soon’: *sceortlicost* [Coaelholm.o3].

Szl ‘better, more effectually, rather, sooner, in preference’: *selest* [Cobenrul.O3, Coboeth.o.02, Coherbar, Colaece.O2, Coverhom], *selost* [Coelive.O3, Colaw1cn.O3, Colaw2cn.O3, Colaw6atr.O3, Colsigewz.O34, Colwsigexa.O35, Contempo.O3, Coquadru.O23], *selust* [coherbar], *soelest* [Codocu2.O12].

Seldor ‘more seldom, less frequently’: *seldost* [Coboeth.o.02].

Sīð ‘late, afterwards’: *siðestan* [Colawine.Ox2], *suiðusð* [CocuraC], *suiðust* [CocuraC]

Smalum ‘little by little’: *smalost* [Colaece.O2].

Smēalice ‘closely, thoroughly, accurately; subtly’: *smeallicost* [Coboeth.o.02].

Sīðlice ‘forcibly’: *siðlucest* [Coneot].

Sweotollice ‘clearly, precisely, plainly, visibly, openly’: *sweotellicost* [Coorosiu.o2], *sweotlost* [Coboeth.o.02], *swiotelecost* [Coverhom], *swiotolusð* [CocuraC].

Swiðe ‘very much, exceedingly, severely, violently, fiercely’: *swiðast* [Coalex.o23, Comarvel.o23], *swiðest* [Coaelholm.o3, cochad.024], *swiðosð* [CocuraC], *swiðost* [cobenrul], *swiðost* [Coaelholm, Coelive.O3, Cobenrul, Coboeth.O.02, Cocathom1, Cocathom2, Cochdrul, Cochronc, Codicts, Coepigen.O3, Coeuphr, Cogenesis, Cogregdc.O24, Coherbar, Colaece.O2, Comarvel.O24, Comary, Coorosiu.O2, Cootest, Coprefcath1.O5, Copreflives, Cosevensl, Coverhoml], *swiðusð* [CocuraC], *swiðust* [coaelholm, cobenrul.o3, Coboeth.o.02, cocathom1, cogregdc.o24, coherbar, Coverhom], *swiðast* [Colacnu.o23], *swiðost* [Coelive.O3, Cobenrul, Coboeth.O.02, Cocathom1, Cochronc, Cogregdc.O24, Coherbar, Colacnu.O23, Colaece.O2, Comart3.O23, Coorosiu.O2, Cowulf.O34], *swiðust* [Cobenrul, Coboeth.O.02, Colaece.O2], *swyðost* [Coaelholm, Coelive.O3, Coalex.O23, Cocathom2, Cochronc, Cogregdh.O23, Coinspolx, Colaw2cn.O3, Comargat, Cosevensl, Cochronc, Coherbar, Coinspold.O34, Colaw6atr.O3, Cootest, Cowulf.O34], *swyðust* [Cobenrul], *swyðast* [Cowulf.o34].

Teartlice ‘sharply, severely (BT)’: *teartlicost* [cobenrul].

Tēonlice ‘in a manner that causes harm or trouble, grievously, miserably (BT)’: *teonlycost* [conicoda].

Tulge ‘strongly, firmly, well’: *tylgest* [Cochad.024].

Dwēorlice ‘insolently’: *ðwyrlicost* [cogregdh.o23].

Ufor ‘higher, further away, further up; later, posterior, subsequent’: *ufemest* [coaelive.o3], *yfemest* [Coboeth.o.02].

Undeore ‘cheap’: *undeorest* [cobenrul].

Ungerædlice ‘sharply, roughly, violently’: *ungeredelīcost* [Coboeth.o.02].

Ūt ‘out; without, outside’: *ytemest* [cobenrul, cowsgosp.o3].

Wærlīce ‘truly’: *wærlīcast* [cowulf.o34].

Wel ‘well, abundantly; very, very easily, very much; fully, quite; nearly; indeed, to be sure’: *wel* [cowulf.o34].

Weorðlice ‘with distinction, gloriously; befittingly (BT)’: *wurðlicost* [coapollo.o3].

Wyr ‘worse’: *wierst* [CocuraC], *wyrest* [conicodC].