

THE LEMMATIZATION OF OLD ENGLISH VERBS FROM THE SECOND WEAK CLASS ON A LEXICAL DATABASE

MARTA TÍO SÁENZ
University of La Rioja
marta_ts91@hotmail.com

ABSTRACT. *This article compiles a list of lemmas of the second class weak verbs of Old English by using the latest version of the lexical database Nerthus, which incorporates the texts of the Dictionary of Old English Corpus. Out of all the inflectional endings, the most distinctive have been selected for lemmatization: the infinitive, the inflected infinitive, the present participle, the past participle, the second person present indicative singular, the present indicative plural, the present subjunctive singular, the first and third person of preterite indicative singular, the second person of the preterite indicative singular, the preterite indicative plural and the preterite subjunctive plural. When it is necessary to regularize, normalization is restricted to correspondences based on dialectal and diachronic variation. The analysis turns out a total of 1,064 lemmas of weak verbs from the second class.*

Keywords: Old English, lexical database Nerthus, verbal morphology, lemmatization, normalization.

LA LEMATIZACIÓN DE LOS VERBOS DÉBILES DE LA CLASE II EN INGLÉS ANTIGUO EN UNA BASE DE DATOS LÉXICA

RESUMEN. *El artículo compila una lista de lemas de los verbos de la clase II en inglés antiguo utilizando para ello la última versión de la base de datos léxica Nerthus, que incorpora los textos del Dictionary of Old English Corpus. De las terminaciones flexivas disponibles, se han seleccionado para esta lematización las más relevantes: el infinitivo, el infinitivo flexionado, el participio de presente, el participio de pasado, la segunda persona singular del presente de indicativo, el plural del presente de indicativo, el singular del presente de subjuntivo, la primera y tercera personas del singular del pretérito de indicativo, la segunda persona del singular del pretérito de indicativo, el plural del pretérito de indicativo y el plural del pretérito de subjuntivo. Cuando ha sido necesario regularizar las grafías, la normalización se ha restringido a las correspondencias basadas en variaciones dialectales y diacrónicas. El análisis da como resultado un total de 1.064 lemas de verbos débiles de la clase II.*

Palabras clave: Inglés antiguo, base de datos léxica Nerthus, morfología verbal, lematización, normalización.

Received 10 February 2015

Revised version accepted 15 March 2015

1. NERTHUS. A LEXICAL DATABASE OF OLD ENGLISH

Martín Arista (2013) presented the new organization of the lexical database of the *Nerthus* project in a lecture delivered at the University of Sheffield.¹ In its latest format, the lexical database is no longer based on dictionary forms but on textual forms. In quantitative terms, this means that the number of files increases from 30,000 to 3,000,000. From the quantitative point of view, the new organization provides all textual occurrences of lemmas together with their context and, therefore, allows to carry out not only morphological and lexical analysis, as the previous version of the database, but also semantic and syntactic analysis. Moreover, all textual variants, frequencies and syntactic patterns can be linked to the dictionary files of the previous version of *Nerthus*. The new database, called *The Grid* consists of five relational layouts, including the dictionary database *Nerthus*, the concordance by word to the Dictionary of Old English Corpus (DOEC), the concordance by fragment to the DOEC, the index to the DOEC (called *Norna*) and the reversed index to the DOEC (called *The Mirror*). Due to copyright reasons, *Nerthus* is the only open access resource (www.nerthusproject).

¹ This research has been funded through the project FFI2011-29532.

com). It contains 30,000 files of lemmatized forms, based primarily on Clark Hall and secondarily on Bosworth-Toller and Sweet. The initial headword list has been compiled by Martín Arista *et al.* (2011) and the meaning definitions provided by the standard dictionaries of Old English above mentioned have been synthesized by Martín Arista and Mateo Mendaza (2013). To briefly illustrate the functionalities of the version of *Nerthus* reviewed in this section, it may be pointed out, in the first place, that the database can turn out the number of textual occurrences of a lemma. For example, *siðian* appears 115 times in the texts. In the second place, the database can break down the occurrences by inflectional form. For instance, the verb *wilnian* occurs in the inflectional forms presented in Figure 1:

Inflectional form	Occurrences	Weak verb 2
<i>wilnast</i>	29	<i>wilnian(ge)</i>
<i>wilniað</i>	135	<i>wilnian(ge)</i>
<i>wilnian</i>	32	<i>wilnian(ge)</i>
<i>wilnianne</i>	3	<i>wilnian(ge)</i>
<i>wilniap</i>	4	<i>wilnian(ge)</i>
<i>wilnie</i>	10	<i>wilnian(ge)</i>
<i>wilniende</i>	13	<i>wilnian(ge)</i>
<i>wilnode</i>	96	<i>wilnian(ge)</i>
<i>wilnoden</i>	1	<i>wilnian(ge)</i>
<i>wilnodest</i>	3	<i>wilnian(ge)</i>
<i>wilnodon</i>	16	<i>wilnian(ge)</i>

Figure 1. Inflections and frequency of *wilnian(ge)*.

Thirdly, the formalism used for representing the prefix *ge-* guarantees the direct link to the *ge-*prefixed counterparts of a given simple verb such as *wilnian*, in Figure 2.

Inflectional form	Occurrences	Weak verb 2
<i>gewilnast</i>	17	<i>wilnian(ge)</i>
<i>gewilnian</i>	47	<i>wilnian(ge)</i>
<i>gewilniap</i>	49	<i>wilnian(ge)</i>
<i>gewilnie</i>	3	<i>wilnian(ge)</i>
<i>gewilniende</i>	3	<i>wilnian(ge)</i>
<i>gewilnige</i>	22	<i>wilnian(ge)</i>
<i>gewilnod</i>	3	<i>wilnian(ge)</i>
<i>gewilnode</i>	120	<i>wilnian(ge)</i>
<i>gewilnodest</i>	3	<i>wilnian(ge)</i>
<i>gewilnodon</i>	13	<i>wilnian(ge)</i>

Figure 2. The prefix *ge-* in *wilnian(ge)*.

And, fourthly, a given inflectional form, such as *wilnast* appears in the fragments, whose short titles are based on Mitchell, Ball and Cameron (1975), that Figure 3 includes.

[Æ LS (Basil) 011600 (392)]	[LS 23 (MaryofEgypt) 008300 (271)]
[Alc 14 (Först) 002200 (28)]	[Met 007200 (5.24)]
[And 008500 (277)]	[Solil 1 001000 (4.5)]
[Bo 013100 (7.16.24)]	[Solil 1 026500 (31.3)]
[Bo 034100 (14.31.13)]	[Solil 1 029200 (33.9)]
[Bo 061500 (23.52.12)]	[Solil 1 031700 (35.17)]
[Bo 086400 (32.71.25)]	[Solil 1 032700 (36.13)]
[Comp 1.6 (Hene) 000600 (3.1)]	[Solil 1 041700 (47.6)]
[CP (Cotton) 0v6100 (33.224.12)]	[Solil 1 041800 (47.12)]
[KtPs 001100 (47)]	[Solil 1 045200 (51.6)]
[LibSc 144100 (45.34)]	[Solil 1 045400 (51.8)]
[LS 10 (Guth) 002000 (5.36)]	[Solil 2 001500 (56.14)]
[LS 10.1 (Guth) 010200 (5.35)]	[Solil 2 005500 (60.8)]

Figure 3. Textual witnesses to *wilnast*.

2. AIMS AND RELEVANCE OF RESEARCH

This article deals with Old English verbs from the second weak class. Its aim is to compile a list of verbal lemmas from this morphological class based on the information provided by the version of the lexical database of Old English Nerthus, as reviewed in the previous section. Therefore, the ultimate source of the data for the analysis is the DOEC, which contains all surviving Old English texts, with a total of approximately 3,000 texts and 3 million words.

The class of the verb has been selected for the analysis because, overall, Old English verbs are morphologically more transparent than nouns and adjectives, which practically share the same endings both in the weak and the strong declension. Within verbs, the class of weak verbs, corresponding to the modern *regular* verbs, has been chosen rather than strong verbs, the counterpart of the modern *irregular* verbs. The reason is that the changes that weak verbs undergo in their inflection take place in the suffixal part of the word rather than in the root, as is the case with strong verbs. Consequently, strong verbs are harder to search. Finally, the second class of weak verbs displays fewer ambiguous inflectional endings than the first class, which can be more easily mistaken for strong verbs.

There are several reasons why it is important to gather such a list of verbal lemmata and to file them into a database. In the first place, standard dictionaries of Old English, including *An Anglo-Saxon Dictionary*, *A Concise Anglo-Saxon Dictionary* and *The*

student's Dictionary of Anglo-Saxon are complete although they are not based on an extensive corpus of the language but on the partial list of sources provided by their prefaces or introductions. In the second place, *The Dictionary of Old English* is based on the corpus mentioned above, but is still in progress (the letter G was published in 2008). Thirdly, with the incorporation of the textual occurrences that correspond to each headword, *Neribus* not only multiplies its size by one hundred but also changes in a qualitative way by linking dictionary forms (types) and textual forms (tokens). This, in turn, will allow us to make advances in the morphological analysis of the language and to carry out quantitative studies in textual frequency. Fourthly, the database format has clear advantages over online corpora. A database can be adapted to the specific needs of a particular research. It can be sorted and searched in ways that online corpora cannot. A database facilitates the definition of relations between data that cannot be captured by online corpora. And the database format allows us to use simultaneously the corpus, the concordance and the index of the language of analysis. Finally, this work can be seen as a contribution to the research programme in the morphology and semantics of Old English represented by Martín Arista (2008, 2010a, 2010b, 2011a, 2011b, 2011c, 2012a, 2012b, 2012c, 2013a, 2013b, 2014), Martín Arista *et al.* (2011), Martín Arista and Mateo Mendaza (2013), Martín Arista and Cortés Rodríguez (2014) and Martín Arista and Veá Escarza (2016).

The remainder of this article is organized as follows. Section 3 presents the morphology of the second class of weak verbs in detail. Section 4 describes the methodology of analysis, which comprises lemmatization and normalization. The different inflectional forms as they appear in the texts have to be related to an abstract form or lemma inflected for a conventional form: in the case of verbs, the infinitive. For instance, given a textual form from the corpus like *hopiað*, it is associated with the infinitive *hopian* 'to hope' by means of a process of lemmatization. Quite often, it is necessary to regularize the forms by means of a process of normalization. For example, when we come across a form like *healsie* we relate it to an infinitive like *bālsian* 'to heal'. Section 5 presents the results of the analysis by inflectional form, lemma and normalization pattern. To close this work, Section 6 draws the main conclusions.

3. RELEVANT ASPECTS OF THE INFLECTION OF THE OLD ENGLISH VERB

This section deals with the characteristics of the three subclasses of weak verbs and their specific features in order to identify the most relevant features of the inflection of the verbs of the second class and to compile a list of formally distinctive inflectional endings that can be used as a starting point in the analysis.

Pyles and Algeo (1982: 125) remark that weak verbs “formed their preterites and past participles in the characteristically Germanic way, by the addition of a suffix containing *d* or immediately after consonants, *t*”. In contrast to strong verbs, these forms do not modify the stem of the verb. Hogg and Fulk (2011: 258) also point out that those suffixes were dental consonants with the function of marking the preterite or past tense. Thus, weak verbs added dental consonants rather than using ablaut or reduplication. In this respect, the most accepted theory is that weak verbs developed their preterite forms from a periphrasis. Pyles and Algeo (1982: 125) hold that many weak verbs were originally causative verbs derived from other categories, such as nouns or adjectives, by means of the “addition of a suffix with an *i*-sound that mutated the stem vowel of the word”. Mitchell and Robinson (1993: 46) add that the stem vowel was normally the same throughout all the verbal forms of the paradigm, which reinforces the idea of regularity and that the inflectional endings of strong and weak verbs showed lots of similarities, although they underwent different evolutions.

Weak class 1 is one of the largest groups of verbs of all the verbal classes in Old English, among other reasons as a result of the just mentioned process of causative stem formation. Class 1 of weak verbs is subdivided into two classes, illustrated by the verbs *fremman* ‘to do’ and *hieran* ‘to hear’. The paradigms of these weak verbs are presented in Figure 4, which is based on Mitchell and Robinson (1993: 46):

Infinitive: subclass 1: *fremman* ‘to do’; subclass 2: *hieran* ‘to hear’

Inflected Infinitive: subclass 1: *tō fremmenne*; subclass 2: *tō hierenne*

Present Participle: subclass 1: *fremmende*; subclass 2: *hierenne*

Past Participle: subclass 1: *(ge-)fremed*; subclass 2: *(ge-)nered*

		<u>Present indicative</u>				<u>Present subjunctive</u>	
		Subclass 1	Subclass 2			Subclass 1	Subclass 2
sg.	1.	<i>fremme</i>	<i>hīere</i>	sg.	1.	<i>fremme</i>	<i>hīere</i>
	2.	<i>fremest</i>	<i>hīerst</i>		2.	<i>fremme</i>	<i>hīere</i>
	3.	<i>fremeþ</i>	<i>hīerþ</i>		3.	<i>fremme</i>	<i>hīere</i>
pl.		<i>fremmaþ</i>	<i>hīeraþ</i>	pl.		<i>fremmen</i>	<i>hīeren</i>
		<u>Present indicative</u>				<u>Present subjunctive</u>	
		Subclass 1	Subclass 2			Subclass 1	Subclass 2
sg.	1.	<i>fremede</i>	<i>hīerde</i>	sg.	1.	<i>fremede</i>	<i>hīerde</i>
	2.	<i>fremedest</i>	<i>hīerdest</i>		2.	<i>fremede</i>	<i>hīerde</i>
	3.	<i>fremede</i>	<i>hīerde</i>		3.	<i>fremede</i>	<i>hīerde</i>
pl.		<i>fremedon</i>	<i>hīerdon</i>	pl.		<i>fremeden</i>	<i>hīerden</i>

	<u>Imperative</u>	
	Subclass 1	Subclass 2
sg.	<i>freme</i>	<i>hïer</i>
pl.	<i>fremmaþ</i>	<i>hïeraþ</i>

Figure 4. The paradigm of class 1 weak verbs *fremman* ‘to do’ and *hïeran* ‘to hear’.

A number of weak verbs had no vowel *i* before the dental preterite suffix in Proto-Germanic, with the consequence that they lack umlaut in the Old English preterite and past participle. In addition, their stems all ended in *-l*, as presented in Figure 5, or velar consonant with the alternation of *toʃ* <cc> and *x* <h>, as shown in Figure 6 (Hogg and Fulk 2011: 274):

<i>cwellan</i> ‘to kill’	<i>cwealde</i>	<i>cweald</i>
<i>dwellan</i> ‘to mislead’	<i>dwealde</i>	<i>dweald</i>
<i>stellan</i> ‘to position’	<i>stealde</i>	<i>steald</i>

Figure 5. Stems in *-l*.

<i>cwecc(e)an</i> ‘to vibrate’	<i>cweahte</i>	<i>cweaht</i>
<i>drecc(e)an</i> ‘to afflict’	<i>dreahte</i>	<i>dreaht</i>
<i>recc(e)an</i> ‘to recount’	<i>reahte, rehte</i>	<i>reaht, reht</i>

Figure 6. Stems in velar consonant.

Campbell (1987: 300) remarks that the 2nd and 3rd persons of the singular (present indicative) of class 1 weak verbs are subject to assimilation. The assimilations of consonants are presented in Figure 7, with an instance of each pattern.

<i>-d-st</i> > <i>-tst</i>	<i>fētst</i> (infinitive <i>fēdan</i> ‘to feed’) then <i>-tst</i> > <i>-st</i> , <i>fēst</i>
<i>-þ-st</i> > <i>tst</i>	<i>cýþst, cýtst</i> (infinitive <i>cýþan</i> ‘to proclaim’)
<i>-g-st</i> > <i>-hst</i>	<i>bīhst</i> (infinitive <i>bīegan</i> ‘to bend’)
<i>-ng-st</i> > <i>-ncst</i>	<i>sprenst</i> (infinitive <i>sprengan</i> ‘to scatter’)
<i>-t-þ, -d-þ</i> > <i>-tt</i>	<i>mētst</i> , (infinitive <i>mētan</i> ‘to measure’)
<i>-s-þ</i> > <i>-st</i>	<i>alýst</i> (infinitive <i>alïesan</i> ‘to free’)
<i>-g-þ</i> > <i>-hþ</i>	<i>bīhþ</i> (infinitive <i>bīegan</i> ‘to bend’)
<i>-ng-þ</i> > <i>ncþ</i>	<i>glencþ</i> (infinitive <i>glengan</i> ‘to decorate’)

Figure 7. Assimilation in the 2nd. and 3rd. person of the singular number.

Moving on to the characteristics of the next class, we find class 2 of weak verbs, the one on which this work focuses. Mitchell and Robinson (1993: 49) remark that this class of verbs “present few problems”. As Hogg puts it (2011: 279), the peculiarity of this class of verbs relies on the fact that this was the only group of verbs which kept adding new verbs during the Old English period. The paradigms of the weak verbs *lufian* ‘to love’ (Mitchell and Robinson. 1993: 49-50), identified as ‘subclass 1’, and the verb *lofi(g)an* ‘to praise’ (Hogg and Fulc 2011: 279-280), identified as ‘subclass 2’, are presented in Figure 8 in order to compare their forms:

<u>Infinitive</u> : subclass 1: <i>lufian</i> ‘love’; subclass 2: <i>lofian</i> ‘praise’	
<u>Inflected infinitive</u> : subclass 1: <i>tō lufienne</i> ; subclass 2: <i>tō lofianne</i>	
<u>Present Participle</u> : subclass 1: <i>lufiende</i> ; subclass 2: <i>lofiende</i>	
<u>Past Participle</u> : subclass 1: <i>(ge-)lufod</i> ; subclass 2: <i>lopod</i>	
<u>Present indicative</u>	
<i>Subclass 1</i>	<i>Subclass 2</i>
sg. 1. <i>lufie</i>	<i>lofige</i>
2. <i>lufast</i>	<i>lofast</i>
3. <i>lufaþ</i>	<i>lofað</i>
pl. <i>lufiaþ</i>	<i>lofiað</i>
<u>Present subjunctive</u>	
<i>Subclass 1</i>	<i>Subclass 2</i>
sg. 1. <i>lufie</i>	<i>lofige</i>
2. <i>lufie</i>	<i>lofige</i>
3. <i>lufie</i>	<i>lofige</i>
pl. <i>lufien</i>	<i>lofigen</i>
<u>Preterite indicative</u>	
<i>Subclass 1</i>	<i>Subclass 2</i>
sg. 1. <i>lufode</i>	<i>lofode</i>
2. <i>lufodest</i>	<i>lofodest</i>
3. <i>lufode</i>	<i>lofode</i>
pl. <i>lufodon</i>	<i>lofodon</i>
<u>Preterite subjunctive</u>	
<i>Subclass 1</i>	<i>Subclass 2</i>
sg. 1. <i>lufode</i>	<i>lofode</i>
2. <i>lufode</i>	<i>lofode</i>
3. <i>lufode</i>	<i>lofode</i>
pl. <i>lufoden</i>	<i>lofoden</i>
<u>Imperative</u>	
<i>Subclass 1</i>	<i>Subclass 2</i>
sg. 1. <i>lufa</i>	<i>lofa</i>
pl. 2. <i>lufiað</i>	<i>lofiað</i>

Figure 8. The paradigm of class 2 weak verbs *lufian* ‘to love’ and *lofi(g)an* ‘to praise’.

Although Hogg and Fulc (2011: 280) notice that “the inflexions of weak verbs of class 2 are, with the exceptions discussed below, the same for all stems, regardless of weight”, these verbs also present some peculiarities, such as contracted forms. As a result of the loss of intervocalic *h*, there were two stems within paradigms like *smēagan* ‘to consider’: *smēag-* and *smēa-* (Campbell 1987: 334), illustrated in Figure 9.

<u>Infinitive</u>	<i>smēagan</i>		
<u>Pres. part.</u>	<i>smēagende</i>		
<u>Pass part.</u>	<i>smēad</i>		
<u>Present indicative</u>		<u>Present subjunctive</u>	
sg.	1. <i>smēage</i>	sg.	1. <i>smēage</i>
	2. <i>smēast</i>		2. <i>smēage</i>
	3. <i>smēaþ</i>		3. <i>smēage</i>
pl.	<i>smēagaþ</i>	pl.	<i>smēagen</i>
<u>Preterite indicative</u>		<u>Preterite subjunctive</u>	
sg.	1. <i>smēade</i>	sg.	1. <i>smēade</i>
	2. <i>smēaest</i>		2. <i>smēade</i>
	3. <i>smēade</i>		3. <i>smēade</i>
pl.	<i>smēadon</i>	pl.	<i>smēaden</i>
<u>Imperative</u>			
sg.	<i>smēa</i>		
pl.	<i>smēagaþ</i>		

Figure 9. The contracted class 2 weak verb *smēagan* ‘to consider.’

The last class of weak verbs is class 3. Hogg and Fulk (2011: 289) explain that “verbs of the third weak class in Germanic are in origin structurally parallel to those of the second weak class” and that the only reason why they became a different class is a vocalic alternation in the formation of the stem. There are just four verbs in class 3, *habban* ‘to have’, *libban* ‘to live’, *secg(e)an* ‘to say’ and *hycg(e)an* ‘to think’ (Campbell 1987: 337), whose paradigms can be seen in Figure 10.

<u>Infinitive</u>	<i>habban</i>	<i>libban</i>	<i>secgan</i>	<i>hycgan</i>
<u>Pres. part.</u>	<i>hæbbende</i>	<i>libbende</i>	<i>secgende</i>	<i>hycgende</i>
<u>Past part.</u>	<i>hæfd</i>	<i>lifd</i>	<i>sægd</i>	<i>hogd</i>
<u>Present indicative</u>				
sg.	1. <i>hæbbe</i>	<i>libbe</i>	<i>secge</i>	<i>hycge</i>
	2. <i>hæfst</i>	<i>leofast</i>	<i>sægst</i>	<i>hygst</i>
	3. <i>hæfþ</i>	<i>leofaþ</i>	<i>sægþ</i>	<i>hygþ</i>
pl.	<i>habbaþ</i>	<i>libbaþ</i>	<i>secgaþ</i>	<i>hycgaþ</i>
<u>Present subjunctive</u>				
sg.	<i>hæbbe</i>	<i>libbe</i>	<i>secge</i>	<i>hycge</i>
pl.	<i>hæbben</i>	<i>libben</i>	<i>secgen</i>	<i>hycgen</i>

Preterite indicative

sg.	1. <i>hæfde</i>	<i>lifde</i>	<i>sægde</i>	<i>hogde</i>
	2. <i>hæfdest</i>	<i>lifdest</i>	<i>sægdest</i>	<i>hogdest</i>
	3. <i>hæfde</i>	<i>lifde</i>	<i>sægde</i>	<i>hogde</i>
pl.	<i>hæfdon</i>	<i>lifdon</i>	<i>sægdon</i>	<i>hogdon</i>

Preterite subjunctive

sg.	<i>hæfde</i>	<i>lifde</i>	<i>sægde</i>	<i>hogde</i>
pl.	<i>hæfden</i>	<i>lifden</i>	<i>sægdēn</i>	<i>hogden</i>

Imperative

sg.	<i>bafa</i>	<i>leoƿa</i>	<i>sæge</i>	<i>hyge</i>
pl.	<i>habbaþ</i>	<i>libbaþ</i>	<i>secgaþ</i>	<i>hycgaþ</i>

Figure 10. The paradigms of class 3 weak verbs *habban* ‘to have’, *libban* ‘to live’, *secg(e)an* ‘to say’ and *hycg(e)an* ‘to think’.

4. METHODOLOGY

The analysis consists of two basic tasks, lemmatization and normalization. As Burkhanov (1998) explains, the first thing we should do when organizing the corpus on which a dictionary is built is to lemmatize the textual (inflected) forms found in the corpus. In this particular case, it would be verbal forms from class 2 of weak verbs. In Burkhanov’s (1998: 122) words “the term ‘lemmatization’ is used to refer to the reduction of inflectional word forms to their lemmata, i.e. basic forms, and the elimination of homography (...) [i]n practice, lemmatization involves the assignment of a uniform heading under which elements of the corpora containing the word forms of same lexeme are represented.” In this respect, Atkins and Rundell (2008: 325) point out that the headword “links all the information about one word together in one entry. In it goes the *canonical form* [italics as in the original] of the headword: the singular of nouns, the infinitive of verbs, the uninflected form of adjectives and adverbs, and so on”. Furthermore, as Jackson (2002: 179) puts it, “the criteria for determining what is a headword have important consequences for lexical description as well as for accessibility”.

In order to find the inflected forms of class 2 weak verbs, it is necessary first of all to choose a set of inflectional endings of these verbs that are representative of their morphology and are not found as inflectional endings in any other classes. The inflections of class 2 weak verbs selected for lemmatization are the infinitive (*-ian*), the inflected infinitive (*-iannē*), the present participle (*-iende*), the past participle (*ge-od*), the first person singular of the present indicative (*-ie/ge-ige*) the second person singular of the present indicative (*-ast*), the present indicative plural (*-iað/-iaþ*), the present subjunctive singular (*-ie/ge-ige*), the first/third person singular of the preterite indicative (*-ode*),

the second person singular of the preterite indicative (*-odest*), the preterite indicative plural (*-odon*) and the preterite subjunctive plural (*-oden*). That is, the *-i-* and *-o-*, characteristic of the second class, that present in the inflectional endings are taken as a distinctive feature that allows us to identify the verbal forms under analysis without ambiguity. These forms comprise the singular and the plural number, the finite and non-finite forms of the verb, the indicative and the subjunctive mood and the present and the preterite tense. Last but not least, these forms are also valid for looking for contracted verbs.

The next step of the analysis is to extract the words ending with these inflections from the DOEC. This has not been done by means of the search engine provided by the online corpus but on the lexical database of Old English *Nerthus*, which comprises, as has been remarked in Section 1, a concordance by fragment and by word of the whole corpus, an index with the number occurrences of all the corpus that lists around 187,000 inflectional forms and a 30,000 file database. The database format has a great advantage over the online corpus: it can search the results of previous searches. Thus, the process of lemma assignment advances on the basis of successive searches that refine little by little the results. With query strings like *==*ode*, *==*ian*, *==*iað*, *==*iap*, *==*ie*, *==*ode*, *==*ie* and *==*iende* the database turns out verbal forms such as *hogode*, *hogian*, *hogiað*, *hogiap*, *hogie*, *gebogode*, *gehagie* and *hogiende* respectively. In the process of lemmatization, these inflectional forms are grouped under the basic form or lemma of *hogian(ge)* (2 occurrences). This does not mean that this process is automatic. In the first place, many undesired results are turned out if the query segment is very short or unspecific. This is the reason why the endings *-ige* and *-od* have been searched only in combination with the prefix *ge-*, thus *ge-ige*, *ge-od*. In the second place, manual work is also needed to find forms that deviate from the paradigms provided by grammars, which tend to represent Early West-Saxon.

At this point, some sort of regularization is necessary that accommodates diachronic, dialectal or textual variants to the grammatical model. Normalization is, in fact, a part of the process of lemmatization and consists of the regularization of non-standard spellings. As Sweet (1976: xi) explains it, “it is often necessary to put the word where the user of the dictionary expects to find it. Therefore, when several spellings of a word appear in the texts, it is necessary to opt for one of them in a consistent way”. For instance, inflected forms such as *bersumie* or *gebersumiað* are found under the lemma *biersumian(ge)* (2 occurrences). *A Concise Anglo-Saxon Dictionary* provides an extensive list of the correspondences it uses for the normalization of Old English texts, but this list has not been used as such because it overnormalizes has many circularities. Instead, the only correspondences that have been selected are those identified by Stark (1982) and de la Cruz (1986) as

constituting instances of dialectal or diachronic variation. The dialect of reference is West-Saxon, in which most surviving texts are written.

Finally, the dictionaries have been necessary for assigning vowel length to lemmas because DOEC does not mark vowel length. The following section presents the results of the application of the methodology just described.

5. RESULTS OF THE ANALYSIS

The following inflected forms have been lemmatized by following the normalization patterns provided. Figure 11 presents the normalization based on intradialectal correspondences while Figure 12 presents the interdialectal basis of normalization.

1. < y > ≈ < ie >

Inflected Form	Lemma
<i>bescyrodest</i>	<i>bescierian</i>
<i>forgymeleasast</i>	<i>forgiemeleasian</i>
<i>forgymeleasodon</i>	<i>forgiemeleasian</i>
<i>gebyrsumast</i>	<i>biersumian(ge)</i>
<i>gebyrsumige</i>	<i>biersumian(ge)</i>
<i>gebyrsumod</i>	<i>biersumian(ge)</i>
<i>gebyrsumodest</i>	<i>biersumian(ge)</i>
<i>gebyrsumodon</i>	<i>biersumian(ge)</i>
<i>gesmyrod</i>	<i>smierwan</i>
<i>geyrsod</i>	<i>iersian(ge)</i>
<i>gyrwast</i>	<i>gierwan(ge)</i>
<i>hyrsumast</i>	<i>biersumian</i>

2. < i > ≈ < ie >

Inflected Form	Lemma
<i>gediglodon</i>	<i>dieglan(ge)</i>
<i>giddodest</i>	<i>gieddian</i>

3. < i > ≈ < y >

Inflected Form	Lemma
<i>asindrodest</i>	<i>āsyndran</i>
<i>gebricgod</i>	<i>brycgian</i>
<i>gemartirod</i>	<i>martyrian</i>

<i>gemartirodon</i>	<i>martyrian</i>
<i>geminegod</i>	<i>mynegian(ge)</i>
<i>genihtsumige</i>	<i>nyhtsumian</i>
<i>gesingod</i>	<i>syngian(ge)</i>
<i>gestirod</i>	<i>styrian(ge)</i>
<i>underwirtwælođen</i>	<i>underwyrwtalian</i>

4. < e > ≈ < ea >

Inflected Form	Lemma
<i>berefodon</i>	<i>berēafian</i>
<i>geernod</i>	<i>earnian(ge)</i>
<i>gemercod</i>	<i>mearcian(ge)</i>
<i>yrfewerdast</i>	<i>yrfeweardian</i>

5. VCC >>> VC

Inflected Form	Lemma
<i>forbttast</i>	<i>forbhtian(ge)</i>
<i>geættrod</i>	<i>ætrian(ge)</i>
<i>geættrodon</i>	<i>ætrian(ge)</i>
<i>gebaddod</i>	<i>hādian(ge)</i>
<i>gebwittod</i>	<i>hwitian(ge)</i>
<i>geliffæstast</i>	<i>līfæstan</i>
<i>geliffestast</i>	<i>līfæstan</i>
<i>gemannod</i>	<i>manian(ge)</i>
<i>gemicclige</i>	<i>miclian</i>
<i>gemicclod</i>	<i>miclian(ge)</i>
<i>gemicclodest</i>	<i>miclian</i>
<i>gerihtwissod</i>	<i>rihtwīsian</i>
<i>gesicclod</i>	<i>siclian</i>
<i>gewissod</i>	<i>wissian(ge)</i>
<i>innseglodon</i>	<i>inseglīan(ge)</i>
<i>mannoden</i>	<i>manian(ge)</i>
<i>spellodon</i>	<i>spelīan</i>
<i>weornnodon</i>	<i>weornīan</i>
<i>widdast</i>	<i>wīdian</i>
<i>willnodon</i>	<i>wilnīan</i>

Figure 11. The intradialectal basis of normalization.

1. < e > ≈ < æ >

Inflected Form	Lemma
<i>geclensod</i>	<i>clǣnsian(ge)</i>
<i>arefnodon</i>	<i>āræfnan</i>
<i>gefegnodon</i>	<i>gefægnian</i>
<i>geliffestast</i>	<i>līfæstan</i>

2. < e > ≈ < ie >

Inflected Form	Lemma
<i>forgemeleasoden</i>	<i>forgiemelesian</i>
<i>gedeglodon</i>	<i>dieglan(ge)</i>
<i>gebersumige</i>	<i>hiersumian(ge)</i>

3. < e > ≈ < ēa >

Inflected Form	Lemma
<i>berefodon</i>	<i>berǣfian</i>

4. < e > ≈ < ea >

Inflected Form	Lemma
<i>geernod</i>	<i>earnian(ge)</i>
<i>gemercod</i>	<i>mearcian(ge)</i>
<i>yrfewerdast</i>	<i>yrfewardian</i>

5. < e > ≈ < eo >

Inflected Form	Lemma
<i>sweðerodon</i>	<i>sweoðerian</i>

6. < æ > ≈ < ēa >

Inflected Form	Lemma
<i>bescæwast</i>	<i>bescǣawian</i>
<i>bescæwodon</i>	<i>bescǣawian</i>
<i>forescæwodeſt</i>	<i>forescǣawian(ge)</i>
<i>gescæwige</i>	<i>scǣawian(ge)</i>

7. < æ > ≈ < ea >

Inflected Form	Lemma
<i>gærcodeſt</i>	<i>gearcian(ge)</i>
<i>gærwodeſt</i>	<i>gearwian(ge)</i>

<i>geærndod</i>	<i>earnian(ge)</i>
<i>gegærwige</i>	<i>gearwian(ge)</i>
<i>gegærwod</i>	<i>gearwian(ge)</i>
<i>gemærcod</i>	<i>mearcian(ge)</i>
<i>gemonigfældod</i>	<i>manigfealdian(ge)</i>
<i>monigfældodest</i>	<i>manigfealdian(ge)</i>
<i>yrfwærdast</i>	<i>yrfeweardian</i>

8. < a > ≈ < ea >

Inflected Form	Lemma
<i>gemarcod</i>	<i>mearcian(ge)</i>
<i>gemonifaldod</i>	<i>manigfealdian(ge)</i>
<i>oferscadodest</i>	<i>ofersceadian</i>

9. < eo > ≈ < e >

Inflected Form	Lemma
<i>streowodon</i>	<i>strēwian(ge)</i>

10. < eo > ≈ < ie >

Inflected Form	Lemma
<i>cleopodon</i>	<i>cliepian</i>

11. < io > ≈ < i >

Inflected Form	Lemma
<i>cliopodon</i>	<i>clipian(ge)</i>

Figure 12. The interdialectal basis of normalization.

The following list of lemmas, which result from lemmatizing the inflectional forms given above, correspond to headwords already filed by the lexical database *Nerthus*: *ābarian*, *ābifian*, *ābisgian*, *āblācian*, *āblindan*, *āblysian*, *āborgian*, *ācēapian*, *āclēnsian*, *ācleopian*, *ācofrian*, *ācōlian*, *ācrammian*, *ācunnian*, *ācwacian*, *ācwician*, *ācwylmian*, *ādēadian*, *ādēafian*, *ādīglian*, *ādibtian*, *ādīlegian*, *ādimmian*, *ādlian*, *ādrūgian*, *ādumbian*, *āfæstnian*, *āfættian*, *āfeormian*, *āfeorsian*, *āfercian*, *āforbtian*, *āfrēfran*, *āfremdan*, *āfūlian*, *āfyllian*, *āgālian*, *āgānian*, *āgeolwian*, *āgēomrian*, *āgīemelēasian*, *āgnian*, *ābalsian*, *ābātian*, *ābeardian*, *ābefegian*, *āblēnsian*, *āblēōdrian*, *ābrēofian*, *ābȳran*, *āidlian*, *ālādian*, *ālatian*, *ālāðian*, *ālēfian*, *āliðian*, *āloccian*, *āmānsumian*, *āmearcian*, *āmeldian*, *āmerian*, *āmolsnian*, *āmundian*, *āmȳrðran*,

anbidian, andbidian, andettan, andswarian, andweardian, andwreðian, angsumian, anlician, āpinsian, āplantian, āpluccian, ārcefnan, ārāsian, ārēcelēasian, āredian, ārēodian, ārian, āsānian, āscamian, āscilian, āscirian, āscian, āscortian, āscrēadian, āscrūtnian, āscian, āsēarian, āslacian, āslāwian, āsmēagan, āsmiðian, āsmorian, āspelian, āstīfian, āstigian, āstundian, āstýfecian, āstyrian, āswārcnian, āswārcan, āsweartian, āsyndrian, ātemian, ātēorian, ātimbran, atolian, ātrahtnian, āwācian, āwēcnian, āwanian, āwannian, āwārnian, āwēodian, āwildian, āwindwian, āwlacian, āwōgian, āwundrian, āwyrwtalian, āðenian, āðēostrian, āðolian, āðrēatian, āðrīstian, āðrōwian, āðrysemian, āðylgian, āðynnian, æfestian, æfnian, æfterfolgian, cēmtigian, cērendian, cēswician, cætclifian, cætfeſtnian, cætlātian, cētrian, cætwenian, cēðmian, barian, bāsnian, baðian, bēacnian, bealdian, becēapian, beceorian, beclypian, beddian, bedecian, bedician, beðigian, bedydrian, beðian, beebbian, beefesian, befæstnian, befician, befōtian, begleddian, begnornian, behæpsian, behættian, behamelian, behāwian, behēafðian, bebegian, behelian, behēofian, bebīwian, behōfian, behogian, behorsian, behrēousian, belādian, belistnian, belūtian, benacian, bencian, bēnsian, beorbhtian, bēotian, berēafian, berēnian, besārgian, bescēawian, besīdian, besmiðian, besmocian, bestreowian, beswician, betarian, beðian, beðrīdian, bewacian, beupnian, bewarenian, bewarin, bewewardian, beweddian, bewitian, bewlēcian, bidian, bifian, bisgian, biterian, bladesian, blendan, blētsian, blissian, blōdgian, blōstmian, blyssian, bōcian, bodian, borgian, bōtian, brastlian, breodwian, brīdlian, brocian, brosnian, brycgian, brýcsian, brytnian, bryttian, bytlan, campian, cealfian, ceallian, cēapian, cearcian, ceorian, ceorlian, circian, clāēnsian, cleacian, clifian, clifrian, clipian, cnēowian, cnēowlian, cnucian, cōlian, corflian, cossian, costian, crafian, crammian, crīstnian, cunnian, cwacian, cwānian, cwedelian, cwician, cwiddian, cwylmian, cynehelmian, dagian, dēafian, dēagian, dīcian, dibtnian, dibtan, dīlegian, drohtian, drohtnian, dropian, droppetan, drūgian, druncnian, dunnian, dwēcian, dwelian, dwellan, dwelsian, dwolian, dysgian, ēacian, ēacnian, ealdian, ealgian, ēanian, eardian, eargian, earmian, earnian, ēaðmōdian, edcwician, edlēanian, ednīwian, edstaðelian, edwitan, efenblissian, efenblēodrian, efenlician, efensārgian, efendrōwian, efesian, efnan, egesian, eglan, ēhtan, elcian, ellenwōdian, elnian, endebyrdan, endian, essian, ēðian, fadian, fāmgian, fandian, fægrian, fāēlsian, fæstian, fæstnian, fæðman, fealgian, fēolian, feormian, feorrian, feorsian, fercian, fician, fiscian, fiðercian, fleardian, floterian, folgian, forcostian, fordīlgian, fordimmian, fordrūgian, forðclypian, forðgelōcian, forðian, forðingian, forðlōcian, forðrysmian, forðyldian, forealdian, foreclipian, foreðingian, foregewissian, forelōcian, forescēawian, forestihtian, foretācnian, forewarnian, forewītegian, foreðingian, forgīemelēasian, forglendrian, forbelian, forbergian, forbogian, forbradian, forbtian, formolsnian, formyrðrian, forrēcelēasian, forrotian,

forscamian, forsēarian, forsetnian, forsmorian, forsorgian, forswigian, forsyngian, fortogian, fortrūwian, forwandian, forweornian, forwisnian, forwundian, forðclypian, forðingian, forðolian, forðrysmān, forðyldian, forðylman, forðyrrian, frætwan, framian, frecnian, frēfran, fremdian, fremman, frēogan, frēolsian, fūlian, fullgearwian, fullwunian, fūlian, fultuman, fultrūwian, fulwian, fyllian, fynegian, fyrclian, fyrðran, gadrian, gālian, gānian, geācolmōdian, geaflian, gearcian, gearwian, gebēagian, gebrytsnian, geðyldgian, gedafenian, geðyldgian, geðyldian, geedbiwian, geedwistian, gefægnian, gefcēlsian, gladian, grāpian, gegrīnian, gebolian, gelangian, geliman, gelōmlīcian, gemōdsumian, genyhtsumian, gēomrian, geondeardian, geondscēawian, geonlīcian, gesadian, gesibsumian, gesingalian, gesundfullian, geswefian, geunārian, geunsōðian, geweddian, gewynsumian, geðyldian, geðyldgian, gīdsian, ginian, gnornian, grāpian, grornian, hādian, hālgian, hālsian, hamelian, hangian, hātian, hāwian, hættian, hēafðian, healtian, hēapian, heardian, hefigian, begian, helmian, hēofian, bergian, hīwian, hlāēnian, hlēoðrian, hlīfian, blinian, hnescian, hōfian, hogian, holian, hopian, horsian, bradian, brēowsian, hrepian, hrisian, hwistlian, hwitian, hȳrian, īdlian, iersian, inearðian, ingelaðian, inlaðian, inseglīan, insomnian, lādian, langian, latian, laðian, lēanian, lēasian, līcian, liðian, lōcian, loccian, losian, lufian, lustfullian, lūtian, lufian, magian, manian, manigfealdian, mēnsumian, mērsian, mearcian, meldian, metgian, metsian, mīdlian, midðolian, miltsian, misbysnian, misfesian, mislīcian, mōdigan, molsnian, morgenwacian, mundian, murcian, myndgian, mynegian, nacian, nemnan, nīwian, notian, ofācsian, ofearmian, oferblissian, ofergitolian, oferbelian, oferbergan, oferhogian, ofermodgian, oferrīcsian, oferscēawian, oferseglīan, ofertrahtnian, oflīcian, ofmyrðrian, ofsceamian, ofsetnian, ofsmorian, ofwundrian, ofðystrian, onāfcestnian, onclīfian, onclypian, ondruncnian, oneardian, onfægnian, ongefæstnian, onhātian, onhāwian, onhȳrsumian, onlōcian, onsceortian, onscunian, onstyrian, onwunian, onðenian, onðracian, openian, ortrūwian, oðehtian, oððingian, pinsian, plantian, pluccian, rēafian, rēnian, rēodian, rīcsian, rotian, sacian, sadian, samnian, sārgian, scāētian, scamian, scēawian, scīmian, scīrian, scrēadian, scrūtnian, scunian, scyldian, sēarian, seglan, sēowan, sibbian, sīdian, sīðian, sigorian, simblīan, slāwian, smēagan, smiðian, smocian, sorgian, sōðian, sparian, spelian, stalian, stibtan, stīfian, stīgian, stibtan, strēlian, strēwian, styrian, sugian, sūrian, suāmian, sweartian, sweotolian, swician, swiðrian, sylvian, symblian, syndrian, syngian, tācnian, temian, tēorian, timbran, tīðian, tōclīfian, tōdīhtnian, tōfēsian, togian, tōhaccian, tōliðian, tōlōcian, tōmearcian, tōsyndrian, trahtnian, trahtian, trēowsian, trucian, trumian, trymian, twiccian, ðancian, ðānian, ðeahtian, ðēofian, ðēostrian, ðēowian, ðingian, ðolian, ðracian, ðrēatian, ðreodian, ðrōwian, ðreodian, ðrīstian, ðrōwian, ðurbhorian, ðurbwunian, ðwēorian, ðynnian, unārwurðian, unclāēnsian,

undercrammian, underplantian, underwreðian, uninseglian, unmynegian, unrōtsian, unsyngian, untrumian, unweorðian, wacian, wānian, wandian, wansian, warian, warnian, wāpnian, wealwian, weardian, welgelician, wellician, wenian, wēodian, weornian, weorðian, wīdlian, wilnian, wītegian, wiðcostian, wiððingian, wiðerbrocian, wiðerweardian, wiðheardian, wiðhogian, wiððingian, wilnian, windwian, wīnbrēafetian, wīsian, wisnian, wistfullian, wītegian, wiðersacian, wiðerweardian, wlacian, wlancian, wlātian, woffian, wōgian, wræcsīðian, wreðian, wrīdian, wrixlan, wuldorfullian, wuldrian, wunian, wundian, wundrian, wunian, wyrwtalian, yfelsacian, ymbðeabtian, ymbfrætewian, ymbhogian, yrfeweardian, yðan.

The following list of lemmas, which result from lemmatizing the inflectional forms given above, cannot be found in the list of reference of the lexical database *Nerthus*. Therefore, they are proposed on the basis of the textual evidence that has been gathered by checking the dictionaries by Sweet and Clark Hall: *ascian, āclian, afandian, anbidian, ārweorðian, āsyndran, ātēorian, aðracian, æmettigian, ændan, æðelian, æwnian, batian, bēagian, besiwan, besorgian, bismarian, bicnan, bicnian, blātian, blisian, bōgian, bolstrian, brehtnian, buterian, byrian, cierran, clānsian, cleofian, cliepien, cræftigian, cwēman, cwician, dafenian, defran, delgian, dēman, dēoran, derian, dolgian, drēfan, drȳgan, dȳrsian, eabtian, earwian, ēastrian, efsian, ēowian, erian, fāgian, fagnian, fabnian, fangian, fægnian, fægenian, fættian, frætwan, feterian, ferian, fetian, findan, fixian, fremian, frēoðan, feðrian, friðian, fullian, fundian, fȳrsian, gaderian, gēomrian, geornian, gieddian, gifian, gītsian, glitenian, gōdian, godspellian, gremian, grennian, grīnian, gristbitian, bagian, hāligian, hæftnian, hælān, hēafian, healdan, healgian, hendan, hēowan, herian, hiersumian, hīewian, blynsian, bnappian, hneppian, bnexian, hrenian, huntian, impian, innian, lācian, lācnian, lafian, lāstfullian, lāswian, leabtrian, lēcnian, leornian, liccian, lichamian, lifæstan, liffæstian, lifæstnian, lifian, lignan, liðigian, leornian, leoðian, lofian, lōgian, luncian, lyfian, macian, mangian, martyrian, mārēian, mæssian, medemian, metan, mettian, micelian, miclian, mōdgian, mōfian, mōtian, murcnian, mynetian, namian, nēadian, nealæcian, nearwian, nēodian, neosian, nerian, niðerian, nipan, nyrwan, nyttian, ofaxian, offrian, ofrian, ofstician, pīlian, platian, plegian, prician, radian, rēadian, rēafian, recenian, regnian, renian, rēonian, reordian, rēstan, rīhsian, ribtan, ribtwisian, ripian, rīsan, rīxian, roscian, rīnian, sadelian, sārēgan, sceorian, scēotan, scotian, segnian, selian, sengan, sēnian, seofan, seofian, seomian, sibsumian, siclian, sigefæstan, singan, singian, slacian, smerian, socian, spyrian, staðolian, stæððan, stician, strangian, stycian, sundrian, sundfullian, suwian, swæðorian, swefan, swefnian, sweogian, swerian, swīgian, swiðian, swornian, sȳclian,*

synnian, syrwan, sȳwian, taccian, talian, tawian, telgian, temprian, teochian, teobian, teobbian, tēoðian, tigðian, tibian, tilian, tīman, tīmian, tintegrian, tōlysan, torfian, tostibtan, tōtian, trēowian, trīwian, tucian, tȳdran, ðafian, ðaccian, ðeafian, ðearfian, ðēawian, ðenian, ðegnian, ðeowtian, ðēodan, ðrafian, ðūhsian, ðwærian, ðyldian, ungeðwærian, untreowsian, ūtlagian, unðwærian, wagian, wædlian, wægnian, wæterian, wearmian, weddian, wemman, wērgian, werian, wician, widmæran, wīfian, winian, wissian, witian, witnian, wiðerian, wracian, wracnian, wuldorbēagian, wynsumian, wyrrian, wȳrtian, wȳrðan, yfelian, yflian, yðgian, yðegan.

6. CONCLUSION

This article has compiled a list of lemmas of the second class weak verbs of Old English by using the latest version of the lexical database *Nerthus*, which incorporates the texts of the *DOEC*. Since this is the beginning of the lemmatization task of the *Nerthus* Project, the most transparent morphological class has been chosen for the analysis, the class 2 weak verb. Out of all the inflectional endings, the most distinctive have been selected for lemmatization: the infinitive (-ian), the inflected infinitive (-ianne), the present participle (-iende), the past participle (ge-od), the second person present indicative singular (-ast), the present indicative plural (-iað/-iap), the present subjunctive singular (-ie/ge-ige), the first and third person of preterite indicative singular (-ode), the second person of the preterite indicative singular (-odest), the preterite indicative plural (-odon) and the preterite subjunctive plural (-oden). A total of 187,000 inflectional forms have been searched for these endings. The searches have been launched on the lexical database of Old English *Nerthus*, which has also filed the results of this analysis and provided a reference list of class 2 weak verbs extracted from its 30,000 word list of lexemes. When it has been necessary to regularize, normalization has been restricted to a number of correspondences based on dialectal and diachronic variation.

A total of 1,064 lemmas of weak verbs from the second class have been found, of which 285 were not on the reference list of *Nerthus*. Since *Nerthus* is based on the standard dictionaries of Old English and provides the information of the dictionary by Clark Hall on an exhaustive way, it seems reasonable to draw the conclusion that after this analysis we have a more accurate knowledge of the relationship between Old English texts and the dictionaries of the language as regards the second class of weak verbs. Moreover, but for *The Dictionary of Old English*, dictionary entries do not contain inflectional forms. Given that *The Dictionary of Old English* has published until the letter G only, the analysis of the letters H-Y that has been carried

out in this work may be seen as a contribution to the field. Apart from proposing lemmas, this work has also helped to improve the information on some lemmas that already appear in dictionaries. This is the case with verbs to which, given the textual evidence, it is necessary to add the prefix *ge-*, as, for instance, *āmerian*, *blýssian*, *cwylmian*, *duwelian*, *fynegian*, *langian*, etc.

REFERENCES

- Atkins, B. T. and M. Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- 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.
- Campbell, A. 1987. *Old English Grammar*. Oxford: Oxford University Press.
- Clark Hall, J. R. 1996 (1896). *A Concise Anglo-Saxon Dictionary*. Toronto: University of Toronto Press.
- De la Cruz, J. 1982. *Iniciación Práctica al Inglés Antiguo*. Madrid: Editorial Alhambra.
- Healey, A. diPaolo (ed.) 2008. *The Dictionary of Old English in Electronic Form A-G*. Toronto: Dictionary of Old English Project, Centre for Medieval Studies, University of Toronto.
- Healey, A. diPaolo (ed.) with J. Price Wilkin and X. Xiang. 2004. *The Dictionary of Old English Web Corpus*. Toronto: Dictionary of Old English Project, Centre for Medieval Studies, University of Toronto. Available online at <http://www.doe.utoronto.ca/pages/pub/web-corpus.html>.
- Hogg, R. M. and R. Fulk. 2011. *A grammar of Old English. Volume 2. Morphology*. Oxford: Blackwell.
- Jackson, H. 2002. *Lexicography. An Introduction*. London: Routledge.
- Lass, R. and J. Anderson. 1975. *Old English Phonology*. Cambridge: Cambridge University Press.
- Martín Arista, J. 2008. "Unification and separation in a functional theory of morphology". *Investigations of the Syntax-Semantics-Pragmatics Interface*. Ed. R. Van Valin. Amsterdam: John Benjamins. 119-145.
- Martín Arista, J. 2010a. "OE strong verbs derived from strong verbs". *SKASE Journal of Theoretical Linguistics* 7: 36-56.
- Martín Arista, J. 2010b. "Lexical negation in Old English". *NOWELE (North-Western European Language Evolution)* 60/61: 89-108.

- Martín Arista, J. 2011a. "Projections and Constructions in Functional Morphology. The Case of Old English *HR OW*". *Language and Linguistics* 12 (2): 393-425.
- Martín Arista, J. 2011b. "Adjective formation and lexical layers in Old English". *English Studies* 92 (3): 323-334.
- Martín Arista, J. 2011c. "Morphological relatedness and zero alternation in Old English". *Morphosyntactic Alternations in English*. Ed. P. Guerrero Medina. Sheffield/Oakville: Equinox. 339-362.
- Martín Arista, J. 2012a. "Lexical database, derivational map and 3D representation". *RESLA-Revista Española de Lingüística Aplicada* (Extra 1): 119-144.
- Martín Arista, J. 2012b. "The Old English Prefix *ge-*: A Panchronic Reappraisal". *Australian Journal of Linguistics* 32 (4): 411-433.
- Martín Arista, J. 2012c. "Nerthus. The reference list of Old English strong verbs". *Working Papers in Early English Lexicology and Lexicography* 2.
- Martín Arista, J. 2013a. "Recursivity, derivational depth and the search for Old English lexical primes". *Studia Neophilologica* 85 (1): 1-21.
- Martín Arista, J. 2013b. Nerthus. Lexical Database of Old English: From Word-Formation to Meaning Construction. Lecture delivered at the English Linguistics Research Seminar (Centre for Research in Humanities), University of Sheffield.
- Martín Arista, J. 2014. "Noun layers in Old English. Mismatches and asymmetry in lexical derivation". *Nordic Journal of English Studies* 13 (3): 160-187.
- Martín Arista, J., E. González Torres, G. Maíz Villalta, R. Mateo Mendaza, C. Novo Urraca, R. Veá Escarza and R. Torre Alonso. 2011. "Nerthus. A lexical database of Old English. The initial headword list 2007-2009". *Working Papers in Early English Lexicology and Lexicography* 1.
- Martín Arista, J. and R. Mateo Mendaza. 2013. "Nerthus. Outline of a lexicon of Old English". *Working Papers in Early English Lexicology and Lexicography* 3.
- Martín Arista, J. and F. Cortés Rodríguez. 2014. "From directionals to telics: meaning construction, word-formation and grammaticalization in Role and Reference Grammar". Eds. M. A. Gómez González, F. Ruiz de Mendoza Ibáñez and F. González García. *Theory and Practice in Functional-Cognitive Space*. Amsterdam: John Benjamins. 229-250.
- Martín Arista, J. and R. Veá Escarza. 2016. "Assessing the semantic transparency of Old English affixation: adjective and noun formation". *English Studies*. 91(1): 61-77.
- Mitchell, B., C. Ball and A. Cameron. 1975. "Short titles of Old English texts". *Anglo-Saxon England* 4: 207-221.
- Mitchell, B. and F. Robinson. 1993. *A guide to Old English*. Cambridge: Blackwell.
- Prokosch, E. 1939. *A Comparative Germanic Grammar*. Philadelphia: University of Pennsylvania.

- Pyles, T. and J. Algeo. 1982. *The origins and development of the English language*. New York: Harcourt Brace Jovanovich.
- Ringe, D. 2006. *From Proto-Indo-European to Proto-Germanic. A Linguistic History of English. Volume I*. Oxford: Oxford University Press.
- Robinson, O. W. 1993. *Old English and its closest relatives: a survey of the earliest Germanic languages*. London: Routledge.
- Smith, J. J. 2009. *Old English: a linguistic introduction*. Cambridge: Cambridge University Press.
- Stark, D. 1982. *The Old English weak verbs. A diachronic and synchronic analysis*. Tübingen: Niemeyer.
- Sweet, M. 1893. "The Third Class of Weak Verbs in Primitive Teutonic, with Special Reference to Its Development in Anglo-Saxon". *The American Journal of Philology* 14 (4): 409-455.
- Sweet, H. 1976 (1896). *The student's Dictionary of Anglo-Saxon*. Cambridge: Cambridge University Press.
- Nerthus*. *A lexical database of Old English* [www.nerthusproject.com].