

1. Introduction

This article presents the theoretical and methodological foundations of the lexical database of Old English *Nerthus* (<http://www.nerthusproject.com>) and underlines the combination of textual and lexicographical corpora required for the design and compilation of this database. By analysing information retrieved from *Nerthus*, two questions related to the general structure of the Old English lexicon are addressed. Firstly, the lexicon is searched for the existence of lexical layers, which link up the notions of derivational process, typological shift and recursivity, while yielding the main building blocks of Old English word-formation. Secondly, the derivational paradigms of strong verbs are quantified, in such a way that the role of strong verbs in lexical derivation is assessed by word class.

2. The *Nerthus* Project: aims and principles

The aim of the *Nerthus* Project is twofold. On the side of lexicology, an exhaustive study of the Old English lexicon in general and word-formation in particular is still pending. In spite of the availability of some insightful analyses, such as Pilch (1970), Kastovsky (1992) and Lass (1994), a comprehensive study of the units, processes and functions of Old English word-formation has still to be written. Moreover, the applications and implications of the study of a lexicon with fairly regular and explicit relations for a general theory of lexical semantics and derivational morphology remain largely unexplored. On the lexicographical side, the most frequently used dictionaries of Old English, including *An Anglo-Saxon Dictionary* (Bosworth and Toller 1973), *The Student's Dictionary of Anglo-Saxon* (Sweet 1976) and *A Concise Anglo-Saxon Dictionary* (Clark Hall 1996), are hardly compatible with current lexicographical practice, while *The Dictionary of Old English* (Healey 2003), which certainly incorporates the latest lexicographical and lexicological advances, has just reached the letter G. Against this background, the tasks ahead of us include to carry out a description of the lexicon of Old English which is based on up-to-date linguistic theory and to publish a product that meets 21st century lexicographical standards. Regarding the theoretical task, the description of the lexicon of Old English will expand further into syntax by relating meaning definitions to logical structures or underlying representations of linguistic expressions containing lexical and syntactic information, as proposed by Role and Reference Grammar (Van Valin and LaPolla 1997; Van Valin 2005). As for the applied task, the exhaustive description of the lexicon is being implemented on a lexical database, which, for the sake of dissemination, can be accessed online by a web browser.

Given these aims, two explanatory principles guide the research.¹ The Principle of Lexical Proto-Grammar stipulates that word-formation units constitute a lexical proto-grammar from which significant syntactic, semantic and pragmatic generalizations can be made. This principle, which draws on the structural-functional tradition of linguistics, as represented by Dik (1997a, 1997b), Foley and Van Valin

* This research has been funded through the project FFI2008-04448/FILO.

¹ The first versions of the methodological and theoretical underpinnings of the lexical database *Nerthus* were presented in Caballero González *et al.* (2004-2005) and Torre Alonso *et al.* (2008).

(1984), Hengeveld and Mackenzie (2008), Van Valin and LaPolla (1997) and Van Valin (2005), constitutes the main theoretical thrust of the project, with far-reaching implications for the development of a theory of derivational morphology (Martín Arista 2008, 2009). For instance, there is a class of weak verbs that derive from adjectives which hold a stative-non stative (ingressive) alternation, including *dimmian* ‘to be or become dim’, *fūlian* ‘to be or become foul’, *heardian* ‘to be or become hard’, *hāsian* ‘to be or become hoarse’, etc. The logical structures of these verbs contain the adjectival base of derivation, as is the case with *dim* ‘dim’ and *dimmian* ‘to be or become dimm’.²

The second explanatory principle that guides the research in the *Nerthus* Project is the Principle of the Targets of Derivation, which establishes that the units of description are the targets rather than the sources of derivation. This is tantamount to saying that predicates are provided with their derivational history, including both the productive and the unproductive patterns. In general, derivation is stepwise, in such a way that each morphological process (affixation, compounding and zero derivation) attaches a maximum of one element to the base. For instance, the strong verb (class IIIb) *ābelgan* ‘to make angry’ has two zero derivatives, the weak verb (class 1) *ābylgan* ‘to irritate’ and the noun *ābylg* ‘anger’, which, in turn, are the bases of derivation for the suffixal nouns *ābylgnes* ‘offence’ and *ābylgð* ‘anger’. The noun *ābylga* ‘anger’ is a derivative by inflectional means from the weak verb while the noun *ābolgennes* ‘irritation’ derives from the past participle of the strong verb *ābelgan* (*ābolgen*). These data are accounted for by the database as follows. Firstly, the entry for *ābelgan* shows that this predicate results for the prefixation of *ā-* to the strong verb (class IIIb) (*ge*)*belgan* ‘to be or become angry’. Secondly, the entry for *ābylgnes* indicates that this noun is produced by the recursive suffixation of *-nes* to the zero derivative *ābylgan*. And, thirdly, the entry for *ābelgan* displays the stem formations of the strong verb, which comprise the form *ābolgen*, on which *ābolgennes* is formed. Given that the suffixation by means of *-nes* is very productive in Old English whereas the prefixation with *a-* is no longer transparent and stem formations are certainly unproductive, the derivational paradigm of BELGAN, contains all the information. Derivational paradigms (as proposed by Pounder 2000), which list all the predicates that can be related morphologically to a given base, have been gathered in order to provide an explanatory account that considers both synchronic productivity and diachronic recoverability (Stark 1982; Kastovsky 1992), as well as to include relevant data even though they do not meet the requirements of either synchronic or diachronic analysis. For example, the derivational paradigm of the strong verb CĪFAN includes *beaducāf* ‘bold in battle’, *cāf* ‘quick; strong; bold’, *cāfe* ‘quickly’, *cāflīce* ‘promptly, vigorously; boldly’, *cāflīc* ‘bold’, *cāfnes* ‘energy’, *cāfscīpe* ‘boldness’, *cīfan* ‘to quarrel’ and *uncāfscīpe* ‘neglect’. As in Pounder (2000), a distinction is made between the lexical paradigm (or product of word-formation processes) and the morphological paradigm (or set of units, processes, principles and rules that produce the lexical paradigm).

In a strictly synchronic analysis, the derivation from a given strong verb proceeds through a weak verb and, more typically, a noun (often neuter), or, less typically an adjective. In the second derivational wave, an adjective is zero-derived from the noun or, less often, the noun is zero-derived from the adjective. For example, the strong verb *þurfan* ‘to need, be required’ is the base of zero derivation for the weak verb *þorfan* ‘to need’ and the noun *þearf* ‘need’, which, in turn, produces the adjective *þearf* ‘needy’, also by means of zero derivation. In diachronic analysis, zero derivation

² See García García (2005) on the formation of causatives in Germanic.

is the counterpart of fully explicit affixation at an earlier linguistic stage.³ Thus, Seebold (1970:509) lists the Germanic derived forms *þarb-a* (adjective), *þarb-ō* (feminine noun), *þarb-ōn* (masculine noun) as corresponding, respectively, to Old English *þearf* ‘needy’, *þearf* ‘need’ and *þearfa* ‘needy person’.⁴

2. From the textual and lexicographical corpora to the lexical database

The principles outlined in the previous section have been implemented on data drawn from a textual corpus as well as a lexicographical corpus. The textual corpus selected for the undertaking is *The Dictionary of Old English Corpus* (<http://www.doe.utoronto.ca/pub/webcorpus.html>), which constitutes the corpus of reference in the field of Old English studies. It gathers 3,047 texts containing approximately 3 million words. The index of the whole corpus consists of 212,231 different words, of which nearly one half occur only once (103,383 is the exact figure). At this stage of the project, the textual corpus has been used mainly for the following tasks:

(i) The completion of the gradual analysis of derivation in order to identify derivational gaps. For instance, given the compound *glīwhlēoðriendlic* ‘musical’, the adjunct *glīw* is listed by dictionaries, whereas the base *hlēoðriendlic* is not. Textual analysis confirms that the attested *hlēoðriend* can be derived from the also attested form *hlēoðrian*, while *hlēoðriendlic* does not appear in the corpus.

(ii) The identification of unlikely derivations that probably constitute calques from Latin. For example, the form *to-for-an-settan* ‘set before’ (Clark Hall 1996) represents an unlikely derivation because it displays three prefixes, which is banned in Old English. Although it is questionable whether *foran* in *toforansettan* represents two prefixes or just one, corpus analysis shows that there are two occurrences only of *to-for-an-settan* in *The Dictionary of Old English Corpus*, both of which appear in Latin-Old English glosses, which suggests that this form is likely to result from a literal translation.

(1)

a. [LibSc 030500 (7.44)]

Alibi dicitur quotiens orantes non cito exaudimur nostra nobis facta in oculis proponamus.

Elles hwar ys gesæd swa oft gebiddende na raðe beoð gehyrede ure us dæda on eagam we **toforansettan**.

b. [RegCGI (Kornexl) 044500 (57.1392)]

Et si fratrem inuenerit somno oppressum, anteponat illi laternam et reuertatur.

& gif he broðor fint mid slæpe ofsetne **toforansette** him þæt leohtfæt & cyrre ongean.

(iii) The isolation of reconstructed forms. Certain forms might have found their way into the dictionaries from glossaries or texts that were not included in the corpus. For instance, the strong verb *cīfan* ‘to quarrel’, which is listed by Bosworth-Toller (1973) but marked as questionable, has no textual occurrences in the corpus even though this dictionary provides the inflective forms *cāf*, *cīfon* and *cīfen*. Although no evidence for

³ See Kastovsky (1968, 1992, 2005) and Marchand (1969) on zero derivation.

⁴ On the evolution of Germanic inflection into Old English, see Bammesberger (1992).

the strong verb can be gathered from the corpus, textual occurrences are found of the derivatives of the strong verb, including *cāf* ‘quick’, *cāfe* ‘quickly’, *cāflīc* ‘bold’, *cāfnes* ‘energy’, *cāflīce* ‘boldly’, *cāfscype* ‘boldness’ and *uncāfscipe* ‘neglect’. It is relatively frequent, in this respect, that forms cannot be found as simplex words in the texts, but they do appear in complex words resulting from compounding and affixation.

(iv) The search for other unattested forms included as headwords by lexicographers. This is the case with *wealdweaxe* ‘sinew’ (Clark Hall 1996), for which no textual evidence is found, except for the expression *þe on wealde weaxeþ* ‘that grows on wood’ [Lch I (HerbHead) 011000 (14.1)], which does not justify the existence of a nominal compound.

(v) The identification of bases of derivation. A very common issue is that lexicographers do not provide the bases of lexical items that are undoubtedly derived. This is the case with adjectives derived by means of *-isc*. Occasionally, the base and the derivative can be found as headwords in the dictionaries, as in *Mierce* ‘Mercia’ and *Miercisc* ‘Mercian’, but more often than not the *-isc* suffixal adjective only is included in the headword list, thus *Indisc* ‘Indian’, *Persisc* ‘Persian’, *Syrisc* ‘Syrian’, etc. The following lines provide evidence for the basic form of these adjectives, respectively:

(2)

a. [Alex 020400 (29.1)]

Ða hit þa on morgen dæg wæs, þa ferdon we on oþer þeodlond **India**, ða cwoman we on sumne micelne feld.

b. [LS 8 (Eust) 006300 (151)]

On þam dagum gelamp þæt eall folc wurþodon symbelnysse mid þam casere þurh þone sige þe he on **Persia** ðeoda gefeaht.

c. [ÆCHom I, 27 011200 (408.228)]

Se cnapa folgode þam mæran witegan eliseum. þa com him to sum rice mann of þam leodscipe þe is **syria** gehaten. his nama wæs naaman & he wæs reoflig: þa becom he to þam godes witegan eliseum on iudea lande.

(vi) The resolution of discrepancies between lexicographers. For instance, Clark Hall (1996) does not include the headword *teartlice* ‘sharply’, which appears in Bosworth-Toller (1973). Textual evidence supports the latter dictionary:

(3)

[Æ LS (Maccabees) 003600 (153)]

Þu afindst his mihte ungefyrn on ðe sylfum, hu he þe tintregað **teartlice** on wítum.

(vii) The identification of the vocalism of the bases of derivational paradigms, which represents a special case of lexicographical discrepancy. For instance, Bosworth-Toller (1973) and Clark Hall (1996) opt for the form *þurfan*, whereas Sweet (1976) prefers the infinitive *þearfan* as corresponding to the preterite singular *þrofte* and the preterite plural *þurfon*. Even though most of the derivatives belonging to the paradigm display the glide *-ea-*, thus *þearf* ‘need’, *þearf* ‘necessary’, *þearfa* ‘poor man’, *þearfa* ‘needing’, *geþearfan* ‘to be in want’, *þearfende* ‘poor person’, *þearfende* ‘poor’, *þearfian* ‘to be in need’, etc.; textual analysis leads to the conclusion that the form with *-ea-* is preferred by the noun and the adjective, whereas the verb favours *-u-*:

(4)

- a. [GD 1 (C) 036000 (9.59.1)]
 & þa þa he geseah, þæt þa **pearfan** genoh hæfdon, he het þone cniht stigan of þære wintreddan & beleac þæt winern & asette his agen insegl on þæt loc.
- b. [And 009900 (337)]
 Ne **ðurfan** ge on þa fore frætwe lædan, gold ne seolfor.

While the textual corpus has been used so far to enlarge or refine specific points, the bulk of the information provided by *Nerthus* has been retrieved from the lexicographical corpus. Although substantially re-interpreted and modified, the morphological and lexical analysis offered by *Nerthus* is based mainly on Clark Hall's *A Concise Anglo-Saxon Dictionary* (including *Supplement*) and, secondarily, on Bosworth-Toller's *An Anglo-Saxon Dictionary* (and *Supplement*), Sweet's *The Student's Dictionary of Anglo-Saxon* and *The Dictionary of Old English*. The derivational paradigms of lexical primes draw on Seebold (1970) and Heidermanns (1993).

Corpus work on Old English texts and dictionaries has resulted in the compilation of a lexical database that contains 29,987 headwords currently, which can be classified by category as can be seen in Table 1:

Category	Predicates
Major lexical classes	
Noun	16,690
Adjective	5,785
Verb	5,618
Adverb	1,654
Minor lexical classes	
Adposition	80
Numeral	52
Pronoun	39
Conjunction	38
Interjection	21
Demonstrative/Article	8
Possessive	3
Total	29,987

Table 1: *Nerthus* headwords by category

By initial letter, the headwords of *Nerthus* are distributed as shown by Table 2:

A 1376	Æ 581	B 1853	C 1074	D 634
E 1155	F 2537	G 2629	(GE-) 1457	H 2489
I 370	L 974	M 1268	N 568	O 1288
P 296	R 565	S 2866	T 990	ð 747
U 1790	V 1	W 2224	Y 255	

Table 2: *Nerthus* headwords by initial letter

Headword definition has been governed by a principle of formal maximization requiring that formal differences of morphological relevance are maximized. To this purpose, numbered entries have been devised, on the grounds of different category, different morphological class or different variants, for predicates otherwise equal. For instance, *besēon* 1 'to see, look, look round', is a class V strong verb, whereas *besēon* 2 'to suffuse' belongs to class I. Each headword is provided with category, spelling variants,

translation into Present-day English, inflectional class and derivational paradigm. As illustration, consider the following derived nouns, for each of which its translation, inflectional class and strong verb base is given in figure 1:

<i>ābylgð</i>	‘anger’	f.	BELGAN
<i>āblācnes</i>	‘pallor’	f.	BLĪCAN
<i>āblāwnes</i>	‘inflation’	f.	BLĀWAN
<i>āblinnendnes</i>	‘cessation’	f.	BLINNAN
<i>ābolgennes</i>	‘irritation’	f.	BELGAN
<i>ābrēotnes</i>	‘extermination’	f.	BRĒOTAN
<i>ācdrenc</i>	‘oak-drink’	m.	DRINCAN
<i>ece</i>	‘ache’	m.	ACAN
<i>ācumendlicnes</i>	‘possibility’	f.	CUMAN
<i>ācunning</i>	‘experience’	f.	CUNNAN

Figure 1: Derived nouns by strong verb base

Whereas figure 1 displays the information from the point of view of derivatives, figure 2 focuses on the base of derivation, such as, for example, BERSTAN:

<i>(ge)berstan</i>	verb	strong IIIc	‘to break, burst, fail, fall; escape; break to pieces’
<i>āberstan</i>	verb	strong IIIc	‘to burst out, break out; escape’
<i>byrst 1</i>	noun	m.	‘loss, calamity, injury, damage, defect’
<i>byrstende</i>	adjective		‘rugiens’
<i>byrstful</i>	adjective		‘disastrous’
<i>byrstig</i>	adjective		‘broken, rugged’
<i>feðorbyrste</i>	adjective		‘split into four’
<i>forberstan</i>	verb	strong IIIc	‘to break, burst asunder, vanish, fail’
<i>fullberstan</i>	verb	strong IIIc	‘to burst completely’
<i>geberst</i>	noun	n.	‘bursting’
<i>mūðberstung</i>	noun	f.	‘eruption of the mouth’
<i>oðberstan</i>	verb	strong IIIc	‘to break away, escape’
<i>tēmbyrst</i>	noun	m.	‘failure to secure a voucher’
<i>tōberstan</i>	verb	strong IIIc	‘to burst apart’
<i>tōberstung</i>	noun	f.	‘bursting’
<i>tōborstennes</i>	noun	f.	‘abscess’
<i>ūtāberstan</i>	verb	strong IIIc	‘to burst out, burst forth’
<i>ūtberstan</i>	verb	strong IIIc	‘to burst out, burst forth’
<i>wiðerbersta</i>	noun	m.	‘adversary’

Figure 2: The derivational paradigm of BERSTAN

As can be seen in figure 3, two types of searches can be conducted on the online database available at <http://www.nerthusproject.com>.



Nerthus Project - Database search

[Exact search](#)[Alphabetical search](#)[Fragmentary search](#)[Word list](#)[Project homepage](#)

Word list

Search for:

Special characters: æ ð

æ 1
æ 3
æalā
æbbung
æbebod
æbēc
æblæce
æbod
æboda
æbrecð
æbrucol
æbs
æbylg
æbylga
æbylgnes
æbylgð
æbære
æcelma

Figure 3: Word search on *Nerthus*.

By whole word, the database can be searched by exact word (exact search) or by initial letter (alphabetical search). By part of a word, a given segment can be searched for, either anywhere in the word (fragmentary search) or at the beginning of the word (word list).



Exact search

Search for:

Special characters: æ ð

ðurfan

Predicate category:

verb

Predicate translation:

to need, be required; must, have occasion to; want, be needy; be under an obligation, owe

Inflectional morphology:

strong with weak forms

Inflectional paradigm:

prot. sing. ðroftø, pres. 1st 3rd p. sing. ðearf, pres. plur. ðurfon, pres. subj. ðurfe, ðyrfe

STEM:

ðURFAN

Derivational paradigm:

beðearfan , beðearfende , beðearflic , beðearfod , beðurfan , feorhðearf , firenðearf , geðearfan , geðearfian , hēahðearf , metedearfende , nearoðearf , nīedbeðearf , nīedðearf 1 , nīedðearf 2 , nīedðearflic , nīedðearflice , nīedðearfnes , nytðearflic , oferðearf , oferðearfa , sāwoldearf , ðearf 1 , ðearf 2 , ðearfa 1 , ðearfa 2 , ðearfednes , ðearfend , ðearfende , ðearfendlic , ðearfendlice , ðearfian , ðearflēas ,

Figure 4: The structure of entries to the database.

Each word is linked to the headword entry, which displays the information shown by figure 4, including predicate category, predicate translation, inflectional morphology, inflectional paradigm, lexical prime and derivational paradigm.

3. Old English lexical primes and the structure of the lexicon

Once the theoretical and methodological underpinnings of database design and compilation have been presented, this section takes issue with the overall structure of the Old English lexicon. Two questions are addressed in this respect, namely the existence of lexical layers and the inventory of lexical primes. The analysis of lexical layers allows for the description of the main building blocks of the Old English lexicon, while the identification of lexical primes leads to quantitative and qualitative conclusions on the core of word-formation at this diachronic stage of the language. Given the paramount importance of strong verbs for the organization of layers, building blocks and paradigms, the identification of lexical primes starts from this morphological class.

Beginning with lexical layers, several authors, such as Kastovsky (1992) and Lass (1994), have stressed the relative regularity and predictability of Old English word-formation, which results in the existence of numerous series of derivatives such as *brōðorlīcnes* ‘brotherliness’, *brōðorlīc* ‘brotherly’, *brōðorrāden* ‘fellowship’, *brōðorscipe* ‘brotherhood’, and *brōðorlēas* ‘brotherless’, if the phenomenon is analysed from the perspective of the base of derivation; or *bōgincel* ‘small bough’, *byrðincel* ‘a little burden’, *cofincel* ‘littlechamber’, *pēowincel* ‘little servant’, *liðincel* ‘little joint’, *rāpincel* ‘small rope’, *scipincel* ‘little ship’, *stānincel* ‘little stone’, *sūlincel* ‘small

furrow', *tūnincel* 'small property' and *wilnincel* 'a little female servant', if derivation is considered from the perspective of the affix. In spite of these regularities, previous studies in Old English derivational morphology are syntagmatic. That is, they are concerned with the affixes (Quirk and Wrenn 1994), processes (Kastovsky 1992; Lass 1994) and the status of the units (Kastovsky 1992, 2006) of lexical creation, but have little to say about derivational series like the ones I have just mentioned and, as a result, are unable to offer an overall description of the lexicon of Old English.

In order to bridge this gap, the analysis of Old English word-formation carried out in this work is paradigmatic in the two ways distinguished by Pounder (2000): the morphological paradigm, which represents the productive part of word-formation, displays units, categories and processes, whereas the lexical paradigm gathers and organizes the product of word-formation, thus accounting for the static part of this phenomenon. An advantage of adopting a paradigmatic approach to word-formation is that derivational paradigms provide morphological and lexical information and combine what is synchronically productive and what is diachronically recoverable, to use Stark's (1982) terms. On this question, Kastovsky (1992:30) states that *a full-scale description of Old English word-formation will have to strike a balance between a purely synchronic and a purely historical-etymological approach by also including unproductive patterns, as long as their output is still transparent*. Let us consider some examples of unproductive patterns. There are, to begin with, non-basic words with explicit inflectional morphemes and without explicit derivational morphemes, as is the case with *rīdan* 'to ride' > *rīda* 'rider'. There are other instances in which no inflectional or derivational morpheme can be analysed, as in *bīdan* 'to delay' > *bīd* 'delay'. A special case of the latter type is raised by non-basic words without analysable morphemes but resulting from ablaut, as is the case with *wyrcean* 'to work' > *weorc* 'work'. There also turn out some instances with ablaut and unproductive (and probably unanalyzable) formatives, such as *-m* in *flēon* 'to fly' > *flēam* 'flight'. Finally, there appear numerous pairs in which no formal difference exists between members of two different lexical categories, as in *mēos* 'moss' > *mēos* 'mossy'. If these derivations are excluded from paradigms, the morphological relationship holding between, for example, the noun *rīda* 'rider' and the strong verb *rīdan* 'ride' is overlooked. Therefore, these instances are unified under the heading of *zero derivation*, which accounts for derivation by inflectional means, as in *rīda* 'rider' as well as derivation without morphemes, as in *weorc* 'work', *bīd* 'delay', *flēam* 'flight' and *mēos* 'moss'.⁵ In short, in the approach adopted in this research, zero derivation contributes the unproductive part of derivation whereas the processes of affixation and compounding provide the productive part of word-formation.

This said, a terminological note is in point. The derivational paradigm, including the morphological and the lexical paradigms, revolves around a lexical item, which constitutes the base of all derivatives. Such a lexical item is termed a *lexical prime* in this framework. This term reflects the fact that all members of the lexical paradigm are morphologically related to the lexical prime and, not less importantly, that the lexical prime contributes a core meaning that is kept, with the modifications caused by subsequent word-formation processes, by most members of the paradigm. For instance, given the lexical paradigm of the prime SPRINGAN, prefixed verbs convey a constant meaning 'to move forward' inherited through word-formation from the strong verb *(ge)springan* 'to jump, leap, spring, burst forth, rise; spread, be diffused, grow; want,

⁵ I refer the reader to González Torres (2009, 2010, forthcoming) on the continuity between inflectional and derivational morphology in Old English.

lack’, thus *āspringan* ‘to spring up or forth, break forth, spread; arise, originate, be born; dwindle, diminish, fail, cease’, *ūpāspringan* ‘to spring up, arise’, *ūpspringan* ‘to rise up’, *tōspringan* ‘to spring apart’, *onspringan* ‘to spring forth’, *ætspringan* ‘to rush forth’. Due to diachronic semantic change, however, some less literal meanings appear, such as *āsprungen* ‘dead’, while still others are clearly figurative, like *ūpsprungennes* ‘eclipse’. In sum, lexical primes, in this framework, stand as semantic and morphological pivots around which derivatives are organized. Such organization is both semantic and morphological. On the semantic side, progressive meaning specialization adds or deletes certain senses while largely maintaining core meanings. On the morphological side, morphological relatedness between base and derivative can be explicit or implicit, as in zero derivation.

With respect to the stratification of the Old English lexicon that can be attributed to the operation of word-formation processes, three derivational layers can be found in the lexicon of Old English, namely zero derivation, affixation and recursive affixation. The layer of zero derivation is further divided into affixless derivation from strong verbs and derivation by inflectional means. Thus, the typological shift from stem-formation to word-formation identified by Kastovsky (1986, 1989, 1990, 1992, 2006) is combined with the notion of recursive (with derived base) vs. non-recursive (with underived base) word-formation. The question is illustrated, in figure 5, with respect to the derivational paradigm of *ÞURFAN*. Double dotted lines represent layer boundaries whereas single dotted lines mark separations between sublayers.

ÞURF

þurfan ‘to need’

ÞEARF

þearfan ‘to need’

ÞORF

=====

ZERO DERIVATION I: STRONG VERB STEMS

þearf 1 ‘need’

þearf 2 ‘necessary’

ZERO DERIVATION II: DERIVATION BY INFLECTIONAL MEANS

þearfa 1 ‘needy person’

þearfa 2 ‘poor’

geðearfan ‘to need’

þearfan ‘to need’

þearfian ‘to be in need’

þearfende 1 ‘poor person’

þearfende 2 ‘needy’

=====

AFFIXATION I: NON-RECURSIVE

be-þurfan ‘to need’

be-ðearfan ‘to need’

ofer-ðearf ‘extreme need’

þearf-end ‘poor man’

þearf-lēas ‘without cause’

þearf-lic ‘necessary’

þearf-līce ‘usefully’

þorf-end ‘needy person’

þorf-fæst ‘useful’

þorf-lēas ‘useless’

un-þearf ‘disadvantage’

AFFIXATION II: RECURSIVE

be-þearf-ende ‘needy’ *þorfendness* ‘poverty’
be-þearf-lic ‘necessarily’
be-þearf-od ‘needy’
þearf-ed-nes ‘poverty’
þearf-end-lic ‘needy’
þearf-lēas-e ‘needlessly’
þearf-end-līce ‘poorly’
þearf-lic-nes ‘poverty’
ofer-þearf-a ‘one in extreme need’
un-þearf-es ‘needlessly’

Figure 5: Lexical layers and the derivational paradigm of ÞURFAN.

As can be seen in figure (5), in the first derivational layer we find strong-verb derivatives which can be broken down into suffixal (*þearfa*) and non-suffixal (*þearf*). By stem, *-ea-* derivatives such as *geðearfan* outnumber *-o-* derivatives like *þorfan*. In the second layer, we come across non-recursive derivatives such as *ofer-ðearf* and *be-þearfan*, from, respectively, the inflectionally unmarked *þearf* and the fully inflective form *þearfan*. Also in the second layer derivatives with derived base are found, such as *þearf-ed-nes* (<*þearf-ed*) and *þearf-end-lic* (<*þearf-end*). As in the first layer, the dominant stem is *þearf-*.

Given these derivational layers, the following building blocks can be identified in the lexicon of Old English: 5,685 derived nouns, including 1,517 zero derivatives and 4,168 with explicit derivation (prefixation and suffixation); 3,865 derived adjectives, which can be broken down into 479 zero derived and 3,386 affixal adjectives; and 4,500 derived verbs, which can be classified as 1,513 zero derivatives and 2,987 affixal verbs. The interpretation of these figures necessarily stresses the fact that affixation, including prefixation and suffixation, constitutes a much more productive word-formation process than zero derivation (the exact figures are 3,553 zero derivatives vs. 11,793 prefixed or affixed derivatives). Zero derivation largely represents an older linguistic stage and, as such, rests as a diachronic relic that is being replaced by the output of other derivational processes. There are numerous synonym pairs of a layer 1 zero derivative and a layer 2 suffixal derivative, such as *lēof/lēoffæst* ‘dear’, *nytt/nyttol* ‘useful’, *stam/stamor* ‘stammering’, *wearm/wearmlic* ‘warm’, etc., that evidence that zero derivation is disappearing to the benefit of other processes of word-formation, notably affixation and compounding. Compounding, which certainly qualifies as a building block of the lexicon of Old English, with 9,986 lexical items, is not organized as a layer but rather interacts quite freely with layers 1 and 2, thus instances of compounding with zero derived base such as *tōðece* ‘toothache’ (from *ece* ‘ache’ < *acan* ‘to ache’) and affixation with zero derived base of the type giflic ‘nuptial’ (from *gift* ‘gift’ < *giefan* ‘to give’).⁶

⁶ See Torre Alonso (2009, 2010) on the interaction of Old English word-formation processes. The related question of lexical integrity is tackled by Martín Arista (2008).

The lexical layers and building blocks that have been defined within this framework, leave us with three pending tasks: (i) the identification of the lexical primes from all lexical categories, (ii) the gathering of synchronic-diachronic paradigms, both morphological and lexical and (iii) the isolation of core meanings.⁷

When it comes to determine where the starting point of lexical derivation lies, Seebold (1970) has convincingly shown that the Germanic strong verb, along with its stem formatives, is central to word-formation. In the same vein, Bammesberger (1965) has underlined the role played by strong verbs in the formation of weak verbs. Focusing on Old English, Kastovsky (1992:297), following Hinderling (1967), remarks *that a description of word-formation in the Germanic languages has to take the strong verb as its starting-point*. In effect, if we choose a lexical paradigm like that of the lexical prime DRIFAN in figure 6,

Ādræfan ‘to drive away’, *ādrīfan* ‘to drive’, *bedrīfan* ‘to beat’, *drāf* ‘action of driving’, *dræfend* ‘hunter’, *drīfan* ‘to stir up’, *eftādrīfan* ‘to reject’, *eftfordrīfan* ‘to drive away’, *fodræfan* ‘to compel’, *fodrīfan* ‘to sweep away’, *fodrīfnes* ‘opposition’, *framādrīfan* ‘to remove’, *framādrīfan* ‘to drive away’, *fullgedrīfen* ‘full of wild beasts’, *(ge)dræfan* ‘to drive’, *(ge)drīf* ‘fever’, *(ge)drīfan* ‘to drive’, *gedrīf* ‘a drive’, *indrīfan* ‘to ejaculate’, *oferdrīfan* ‘to overcome’, *onwegādrīfan* ‘to drive away’, *onwegādrīfennes* ‘a driving away’, *tōdræfan* ‘to scatter’, *tōdræfednes* ‘dispersion’, *tōdrīfan* ‘to scatter’, *þurhdrīfan* ‘to drive through’, *ūtadræfan* ‘to drive out’, *ūtadrīfan* ‘to drive out’, *ūtdræf* ‘decree of expulsion’, *ūtdræfere* ‘driver out’, *ūtdrīfan* ‘to expel’, *underdrīfennes* ‘subjection’, *undrīfen* ‘not driven or tossed’, *wiðdrīfan* ‘to repel’

Figure 6: The lexical paradigm of DRIFAN

In order to determine the processes of word-formation that constitute the corresponding morphological paradigm of a set of derivatives like the one in figure 6, a distinction has to be drawn between the diachronic and the synchronic axis. In the diachronic axis, the derivatives with *æ* like *ūtdræf* ‘decree of expulsion’ derive from the Germanic weak verb **draibjanan* > Old English *(ge)dræfan* ‘to drive’ (Holthausen 1963:75; Seebold 1970:163; Orel 2003:74). In the synchronic axis, *dræf* holds a vocalic alternation with the preterite singular form of the strong verb *drāf* of the seventh vocalic type (A7) identified by Kastovsky (1968:67), which is due to *i*-mutation and involves the back vowel *a* and the front vowel *æ*. At the same time, strong verbs such as *eftādrīfan* ‘to reject’ derive from the infinitive of the basic strong verb, while the noun *drāf* ‘action of driving’ derives from the preterite form of the strong verb and the adjective *undrīfen* ‘not driven or tossed’ derives from the past participle. The panchronic analysis for which derivational paradigms allow stresses the paramount importance of strong verbs in Old English word-formation. Indeed, strong verbs not only turn out derivatives belonging to other lexical classes but also constitute the base of derivation of other strong verbs (*drīfan* ‘drive’ > *eftādrīfan* ‘to reject’, *framādrīfan* ‘to drive away’, *þurhdrīfan* ‘to drive through’, etc.), which, in turn, produce new derivations, as in *drīfan* ‘drive’ > *fodrīfan* ‘to sweep away’ > *fodrīfnes* ‘opposition’.

This is not to say, however, that lexical primes invariably originate in strong verbs. In other words, some strong verbs qualify as lexically derived from either other strong verbs or, much more frequently, nouns and adjectives. Leaving aside the

⁷ Although I will not have much to say about core meanings in this work, the expression of semantic universals in Old English has been dealt with by Martín Arista and Martín de la Rosa (2006), de la Cruz Cabanillas (2007) and Guarddon Anelo (2009a, 2009b).

derivation from strong verbs, to which I have just referred, and focusing on denominal strong verbs, Pilch (1970:132) finds six instances of the seventh class, including *rædan* ‘to advise’ (<*ræd* ‘advice’), *slæpan* ‘to sleep’ (<*slæp* ‘sleep’), *blandan* ‘to blend’ (<*gebland* ‘blend’), *hrōpan* ‘to shout’ (<*hrōp* ‘cry’), as well as the mutated *wēpan* ‘to weep’ (<*wōp* ‘weeping’) and *spætan* ‘to spit’ (<*spāt* ‘spittle’). Holthausen (1963:344) gives the noun *teld* ‘tent’ as the source of (*ge*)*teldan* ‘to spread a covering’. Further evidence of denominal formation of strong verbs includes *blōtan* ‘to kill for sacrifice’ (<*blōt* ‘blod’), *edgyldan* ‘to remunerate’ (<*gold* ‘gold’), *oferreccan* ‘to convince’ (<*racu* ‘explanation’) and *ofstānan* ‘to stone’ (<*stān* ‘stone’). Turning to deadjectival strong verbs, Heidermanns (1993:40) relates (*ge*)*brædan* ‘to make broad’ (<*brād*) to the Germanic primary adjective *breida-* ‘broad’. The same author (1993:576) refers us to the adjective **sweiga* ‘still’ for *swōgan* ‘to sound’, although Kastovsky (1968:109) analyses the morphological relation between *swōgan* and the noun *swēg* ‘sound’ and identifies a vocalic alternation (A8) and a consonantal one (C4) caused, respectively, by *i*-mutation and palatalization, which probably reinforces the basis character of the strong verb with respect to the adjective. Further evidence of deadjectival strong verb formation can be gathered, including (*ge*)*hrēran* ‘to move’ (<*hrōr* ‘active’), (*ge*)*sceorpan* ‘to scrape’ (<*scearp* ‘sharp’) and (*ge*)*swīðan* ‘to strengthen’ (<*swīð* ‘strong’).

Overall, *Nerthus* yields 12,525 lexemes that belong to the derivational paradigms of strong verbs, out of a total 29,987 in all the lexicon, which amounts to 41.8% of the total amount of lexical items in the vocabulary of Old English as gathered by the lexical database. By lexical class and process, the whole inventory of derivatives can be broken down as shown by table 3.

	Adjectives	Adverbs	Nouns	Verbs	Total
Compounding	675	30	3,595	118	4,418
Prefixation	571	42	332	1,708	2,653
Suffixation	824	363	1,652	78	2,917
Zero derivation	371	54	1,232	880	2,499
TOTAL	2,441	489	6,811	2,784	12,525

Table 3: The derivational paradigm of strong verbs (lexical).

The figures in table 3 call for some explanation. The processes of word-formation that produce the lexical paradigms comprise, together with zero derivation and compounding, prefixation and suffixation. For instance, the paradigm of CALAN contains, together with the prime *calan* ‘to grow cold’, the zero derived *ciele* ‘cold’ and the compound *færcyle* ‘intense cold’ as well as the prefixed *ācalan* ‘to become frost-bitten’ and *ofcalan* ‘to make or grow cold’ and the suffixed *forcilled* ‘chilled’. The figures in table 3 include all derivatives, both direct and indirect, from strong verbs. In the paradigm of ACAN, for example, the zero derived noun *ece* ‘pain’, which can be directly related to the strong verb, is found along with the compounds *fōtece* ‘gout’, *heortece* ‘heartache’, *bānece* ‘pain in the thigh’, etc., which can be morphologically related to the strong verb indirectly only, that is, through *ece*. Compounds are listed under the two paradigms to which they belong. As illustration, the compounds *bearnēaca* ‘pregnant’, *bearnēacen* ‘pregnant’, *bearnēacnod* ‘pregnant’ and *bearnēacnigende* ‘pregnant’ are displayed by the paradigms of BERAN and ĒACAN.

The figures in table 3 can be interpreted as follows. The total number of derivational paradigms with bases of the strong verb class is 210, which turns out an average paradigm productivity of 59.6 derivatives. By morphological process,

compounding is the most productive, yielding 35.2% of non-basic terms. The other processes produce similar percentages of derivatives, above 20%. By lexical class, strong verbs produce nouns mainly, and verbs, adjectives and adverbs on a smaller scale. These figures are more explanatory when the number of derivatives by lexical category is compared with the number of items of each class. As shown in table 1, there are 16,690 nouns in the lexicon of Old English as described by *Nerthus*, which means that 40.8% of nouns are morphologically related to strong verbs. As for adjectives, 42.1% of the total amount of items of this class qualify as strong verb derivatives. The case with verbs is different. Approximately one half of the total inventory of Old English verbs derive from strong verbs, the other half comprising basically weak verbs with nominal or adjectival base. Adverbs stand at the other extreme. Indeed, 29.5% of adverbs only can be traced back to strong verbs.

4. Concluding remarks

The lexical database of Old English *Nerthus* has been designed and compiled with a view to offering an exhaustive description of the lexicon and, ultimately, an explanation for lexical and morphological phenomena based on functional principles. Although *Nerthus* represents still research in progress and, as such, is subject to progressive modification, enlargement and refinement, it has already proved a superb tool for determining the overall structure of the Old English lexical stock as well as for investigating the lexical primes of Old English word-formation.

By combining a textual and a lexicographical corpus, the research carried out so far has drawn an overall picture of the Old English lexicon in which strong verbs play a central role and zero derivation loses ground to affixation and compounding. Moreover, some lexical categories, like the verb, stand out as endocentric categories (in the sense of selecting derivation bases of the same lexical class), whereas others, such as the adverb, are clearly exocentric, given that they are derived mainly from members of other lexical categories.

To round off, the progress made in the identification of strong verb lexical primes and the gathering of synchronic-diachronic paradigms advises to apply the methodology to other categories and stresses the necessity of isolating the core meanings of lexical derivation. This task is pending for future research.

References

- Bammesberger, A. 1965. *Deverbative jan-Verba des Altenglischen, vergleichend mit den übrigen altgermanischen Sprachen dargestellt*. München: Ludwig-Maximilians Universität.
- Bammesberger, A. 1992. The place of English in Germanic and Indo-European. In R. Hogg (ed.), *The Cambridge History of the English Language I: The Beginnings to 1066*. Cambridge: Cambridge University Press. 26-66.
- Bosworth, J. and T. N. Toller. 1973 (1898). *An Anglo-Saxon Dictionary*. Oxford: Oxford University Press.
- Caballero González, L., E. González Torres, A. Ibáñez Moreno and J. Martín Arista. 2004-2005. Predicados verbales primitivos y derivados en inglés antiguo. Implicaciones para la elaboración de una base de datos léxica. *Revista Española de Lingüística Aplicada*. 17-18: 35-49.
- Clark Hall, J. R. 1996 (1896). *A Concise Anglo-Saxon Dictionary*. Toronto: University of Toronto Press.

- de la Cruz Cabanillas, I. 2007. Semantic primes in Old English: a preliminary study of descriptors. *SELIM*. 14: 37-58
- Dik, S. 1997a (1989). *The Theory of Functional Grammar. Part 1: The Structure of the Clause*. Edited by K. Hengeveld. Berlin: Mouton de Gruyter.
- Dik, S. 1997b. *The Theory of Functional Grammar, Vol. 2: Complex and Derived Constructions*. Edited by K. Hengeveld. Berlin: Mouton de Gruyter.
- Foley, W. and R. Van Valin. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- García García, L. 2005. *Germanische Kausativbildung*. Göttingen: Vandenhoeck & Ruprecht.
- González Torres, E. 2009. *Affixal Nouns in Old English: Morphological Description, Multiple Bases and Recursivity*. PhD Dissertation, Department of Modern Languages, University of La Rioja.
- González Torres, E. 2010. The Continuum Inflection-Derivation and the Old English suffixes *-a, -e, -o, -u*. *ATLANTIS* 32.1: 103-122.
- González Torres, E. The bases of derivation of Old English affixed nouns: status and category. *Studia Anglica Posnaniensia*. Forthcoming.
- Guarddon Anelo, M. C. 2009a. Un análisis de las propiedades combinatorias de los primitivos semánticos a través de las adposiciones complejas en inglés antiguo. *Revista Española de Lingüística*. Forthcoming.
- Guarddon Anelo, M. C. 2009b. The Natural Semantic Metalanguage of Old English Compound Adpositions. *English Studies*. Forthcoming.
- Healey, A. di Paolo. 2003. *The Dictionary of Old English in Electronic Form A-F*. Toronto: Dictionary of Old English Project, Centre for Medieval Studies, University of Toronto.
- Heidermanns, F. 1993. *Etymologisches Wörterbuch der germanischen Primäradjektive*. Berlin: Walter de Gruyter.
- Hengeveld, K. and J. Lachlan Mackenzie. 2008. *Functional Discourse Grammar. A typologically-based theory of language structure*. Oxford: Oxford University Press.
- Hinderling, R. 1967. *Studien zu den starken Verbalabstrakten des Germanischen*. Berlin: Walter de Gruyter.
- Holthausen, F. 1963. *Altenglisches etymologisches Wörterbuch*. Heidelberg: Carl Winter Universitätsverlag.
- Kastovsky, D. 1968. *Old English Deverbal Substantives Derived by means of a Zero Morpheme*. PhD Dissertation. Tübingen: Eberhard-Karls-Universität.
- Kastovsky, D. 1986. Deverbal nouns in Old and Modern English: from stem-formation to word-formation. In J. Fisiak (ed.), *Historical Semantics-Historical Word Formation*. 221-261. Berlin: Mouton de Gruyter.
- Kastovsky, D. 1989. Typological Changes in the History of English Morphology. In U. Fries and M. Heusser (eds.), *Meaning and Beyond. Ernst Leisi zum 70. Geburtstag*. 281-293. Tübingen.
- Kastovsky, D. 1990. The typological status of Old English Word Formation. In S. Adamson, V. Law, N. Vincent and S. Wright (eds.), *Papers from the 5th International Conference on English Historical Linguistics*. 205-224. Amsterdam: John Benjamins.
- Kastovsky, D. 1992. Semantics and vocabulary. In R. Hogg (ed.), *The Cambridge History of the English Language I: The Beginnings to 1066*. Cambridge: Cambridge University Press. 290-408.

- Kastovsky, D. 2005. Conversion and/on zero: word-formation theory, historical linguistics, and typology. In L. Bauer and S. Varela (eds.) *Approaches to Conversion/Zero Derivation* 31-50. Münster: Waxmann.
- Kastovsky, D. 2006. Typological Changes in Derivational Morphology. In A. van Kemenade and B. Los (eds.), *The Handbook of The History of English*. 151-177. Oxford: Blackwell.
- Lass, R. 1994. *Old English. A historical linguistic companion*. Cambridge: Cambridge University Press.
- Marchand, H. (1969) *The Categories and Types of Present-Day English Word-Formation*. München: C.H. Beck'sche.
- Martín Arista, J. 2008. Unification and separation in a functional theory of morphology. Robert Van Valin (ed.), *Investigations of the Syntax-Semantics-Pragmatics Interface*. Amsterdam: John Benjamins. 119-145.
- Martín Arista, J. 2009. A Typology of Morphological Constructions. In C. Butler and J. Martín Arista (eds.), *Deconstructing Constructions*. Amsterdam: John Benjamins. 85-115.
- Martín Arista, J. and V. Martín de la Rosa. 2006. Old English Semantic Primes: Substantives, Determiners and Quantifiers. *ATLANTIS* 17: 9-28.
- Orel, V. 2003. *A Handbook of Germanic Etymology*. Leiden: Brill.
- Pilch, H. 1970. *Altenglische Grammatik*. Max Hueber Verlag, München.
- Pounder, A. 2000. *Processes and Paradigms in Word-Formation Morphology*. Berlin: Mouton de Gruyter.
- Quirk R. and C. L. Wrenn. 1994 (1955). *An Old English Grammar*. DeKalb, Illinois: Northern Illinois University Press.
- Seebold, E. 1970. *Vergleichendes und etymologisches Wörterbuch der germanischen starken Verben*. The Hague: Mouton.
- Stark, D. 1982. *The Old English weak verbs. A diachronic and synchronic analysis*. Tübingen: Niemeyer.
- Sweet, H. 1976 (1896). *The Student's Dictionary of Anglo-Saxon*. Cambridge: Cambridge University Press.
- Torre Alonso, R. 2009. *Morphological Process Feeding in the Formation of Old English Nouns: Zero Derivation, Affixation and Compounding*. PhD Dissertation, Department of Modern Languages, University of La Rioja.
- Torre Alonso, R. 2010. Morphological process feeding in the formation of Old English nouns. *SKASE Journal of Theoretical Linguistics* 7: 57-69.
- Torre Alonso, R., J. Martín Arista, A. Ibáñez Moreno, E. González Torres and L. Caballero González. 2008. Fundamentos empíricos y metodológicos de una base de datos léxica de la morfología derivativa del inglés antiguo. *Revista de Lingüística y Lenguas Aplicadas* 3: 129-144.
- Van Valin, R. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.
- Van Valin, R. and LaPolla, R. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.