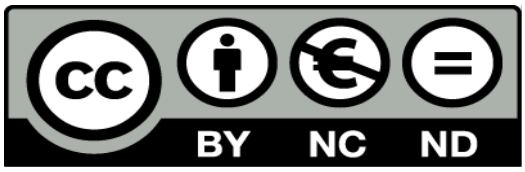
UNIVERSIDAD DE LA RIOJA

TESIS DOCTORAL

Título
Lemmatisation of old english strong verbs on a lexical database
Autor/es
Darío Metola Rodríguez
Director/es
Francisco Javier Martín Arista y Roberto Torre Alonso
Facultad
Facultad de Letras y de la Educación
Titulación
Departamento
Filologías Modernas
Curso Académico
2014-2015



LEMMATISATION OF OLD ENGLISH STRONG VERBS ON A LEXICAL DATABASE

Darío Metola Rodríguez

Tesis doctoral 2015

Departamento de Filologías Modernas

Universidad de La Rioja

LEMMATISATION OF OLD ENGLISH STRONG VERBS ON A LEXICAL DATABASE

Darío Metola Rodríguez

PhD Dissertation supervised by

Dr. Javier Martín Arista

and

Dr. Roberto Torre Alonso

2015

Departamento de Filologías Modernas

Universidad de La Rioja

Acknowledgements

This research has been funded through the projects FFI2011-29532 and FFI2014-59110, which are thankfully acknowledged. Thanks are also due to the University of La Rioja for providing me with the means necessary to carry out this research.

I am very grateful to Javier Martín Arista and Roberto Torre Alonso for their support and help throughout this research and the previous years of work. Their hard work and dedication was essential for the successful end of this research, and without them the result would not have been the same.

I would also like to thank the research group *Nerthus*, and their help and understanding in these last years, in which we have shared very good moments.

Special thanks to Carmen, whose pieces of advice helped me a lot before getting involved in this new adventure, and while I was deeply into it.

CONTENTS

Main abbreviations

acc.: accusative.

adj.: adjective.

adv.: adverb.

att. sp.: attested spellings.

BT.: Bosworth-Toller's *An Ango-Saxon Dictionary*.

cf.: compare.

cl.: class.

conj.: conjunction.

DOE.: *The Dictionary of Old English in Electronic Form A-G*.

Du.: Dutch.

f.: feminine.

fc.: forthcoming.

Fr.: French.

gen.: genitive.

impers.: impersonal.

inf.: infinitive.

intr.: intransitive.

m.: masculine.

mf.: masculine and feminine.

mfn.: masculine, feminine and neuter.

n.: neuter.

nom.: nominative.

occ.: occurrence.

OE.: Old English.

p.: past.

part.: participle.

PDE.: Present Day English.

PIE.: Proto-Indo-European.

pl.: plural.

pp.: past participle.

pres.: present.

pret.: preterite.

pron.: pronoun.

sg.: singular.

str.: strong.

subj.: subjunctive.

s.v.: *sub voce*.

tr/trans.: transitive.

vb /v.: verb.

wk.: weak.

Index of figures

Chapter 2

Chapter 3

Chapter 5

Index of tables

Chapter 3

Chapter 5

1. INTRODUCTION

1.1. Overview

The new database of the Nerthus Project, called *The Grid*, was presented by Martín Arista (2013) in a lecture delivered at the University of Sheffield. *The Grid* consists of five relational layouts, namely *Nerthus* and a concordance by fragment, a concordance by word, an index and a reverse index of the *Dictionary of Old English Corpus*. *The Grid*, therefore, is no longer a based on dictionary forms but on textual attestations. From the quantitative point of view, this change represents an increase from 30,000 database files to 3,000,000. Apart from the quantitative question, the new design of the database has significant advantages over the dictionary database *Nerthus*. Firstly, the current dictionaries of Old English can be related to the occurrences of the lemmatized entries that they contain, which allows researchers to carry out not only morphological but also lexical and syntactic work. Secondly, it is possible to link to the *ge*-prefixed forms to the corresponding simplex verb. Thirdly, all the contexts of a given inflectional form can be accessed in their respective sources. And finally, after launching the relevant queries and checking the results manually, it is possible to gather all the inflectional forms of a variable lexical category.

Of the many lines of research that the new organization of the database opens, this work takes issue with the lemmatization of textual forms. The reason for this choice is that a morphologically annotated corpus of Old English is still a pending task of the discipline. The morphological information is available for the letters A-G, which have already been published by the *Dictionary of Old English*, but it is not present or easily searchable in the current dictionaries for the letters H-Y.

The starting point of this research is that the task of lemmatization can be done partly automatically and partly manually. The information contained in the lexical database together with the functionalities of Filemaker software can maximize the automatic part of the analysis and mimimize manual revision.

The aim of the research is to provide an inventory of lemmas of strong verbs that is based on the textual evidence provided by the *Dictionary of Old English Corpus*. With respect to the Nerthus Project, this thesis aims at developing a search system based on successive searches, in such a way that more transparent forms are tagged before more opaque forms.

The restriction of the scope of the analysis to the strong verbs is based on two reasons. Firstly, the strong verb system in Old English plays a central role in the derivation and development of the lexicon. On the other hand, strong verbs, being characterised by *ablaut* or vowel gradation, can be searched not only by inflectional ending but also by root vowel, which contributes to refining the search system. Furthermore, the lemmatization of the strong verbs paves the way for the search for the other classes, once strong verbs have been identified.

The methodology comprises three main steps: the compilation of a corpus of strong verbs suitable for the analysis, the identification of inflectional forms and the definition of automatised search codes. The list of strong verbs has been taken from Martín Arista (2012), who draws for the seven classes of strong verbs on Campbell (1987) and Hogg and Fulk (2011) and on Krygier (1994) for the subclasses. For the identification of the relevant inflectional forms, the underived strong verbs have inflected for the infinitive, present indicative (singular and plural), preterit indicative (singular and plural), present subjunctive (singular and plural), preterit subjunctive (singular and plural), and imperative (singular and plural). As for the search codes, I have proposed a system of four successive searches codes which are specifically designed to look for particular verb forms.

Apart from the inventory of strong verbs, conclusions are expected in two areas. In the first place, this thesis can answer in a motivated way the question of the limits of automatisation in morphological analysis. Secondly, this thesis can also throw light on the question of the regularisation of forms in lexicographical work or *normalisation*.

1.2 Chapter organisation

Leaving aside the introduction, this doctoral research is organised as follows:

Chapter 2 sets the lexicological basis of this work, as the main aim of this dissertation is to offer a lemmatised inventory of the strong verbs of Old English on the grounds of the *Dictionary of Old English Corpus*. Although the focus of this research lies on the lexicographical side, a solid lexicological grounding in also needed. Thus, section 2.2 offers some general remarks on the linguistic field of lexicology, including both the non-derived and the derived lexicon whereas section 2.3 goes on to discuss aspects proper of Old English lexicology. Section 2.4 is devoted to the description of the origin and evolution of the strong verb system from Indo-European (2.4.1), through Germanic (2.4.2), to Old English (2.4.3). At this stage, a full description of the verb system in Old English and its impact on the language and the lexicon follows. Thus, sub-section 2.4.3.1 concentrates on the different strong verb classifications, whereas 2.4.3.2 focuses on dialectal variation and 2.4.3.3 describes derivation originating in the strong verbs.

Chapter 3 plunges into the role played by lexicography and its evolution from the philological work of earlier times to the current linguistic research and the growth in importance and influence of the new technologies in the compilation, development and distribution of dictionaries. It presents, then, a more applied view of the task to be undertaken. Section 3.2 discusses not only the goals of lexicography, that is, how lexicography focuses its attention on the inclusion of words into dictionaries, and in the correspondence of form and meaning, but also emphasises the changes undergone by the discipline while presenting an updated framework for the current research. Section 3.3 focuses on the stage of the language at stake and provides a contrastive review of the major Old English dictionaries, including *An Anglo-Saxon Dictionary* (Bosworth and Toller 1973), *The student's Dictionary of Anglo-Saxon* (Sweet 1976), *A Concise Anglo-Saxon Dictionary* (Clark Hall 1996), and *The Dictionary of Old English in Electronic Form A-G* (Healey *et al.* 2008). This exhaustive comparison is done at different levels. The first comparison refers to headword spelling in subsection 3.3.1. Subsection 3.3.2. focuses on the scope of the four dictionaries, how the headwords are displayed, and the information that is included for each headword. Subsection 3.3.3. is devoted to alternative spellings and cross-references. Subsection 3.3.4 deals with phonology and discusses issues such as vocalic quantity. Subsection 3.3.5. concentrates on the

syntactic information offered by each dictionary Subsection 3.3.6. faces the question of morphology, both derivational and inflectional. Meaning definitions, senses and semantics in general are compared in Subsection 3.3.7. Subsection 3.3.8. pays attention to etymology, and the differences that can be found in the quantity and quality of the information provided by the dictionaries. Subsection 3.3.9. is devoted to the treatment of textual evidence that supports the inclusion of a given headword or form in the dictionary. After this review of Old English dictionaries, the relevance, goals and scope of the research are presented in section 3.4, while section 3.5 is devoted to the presentation of the methodological guidelines underlining the research, along with the steps followed along the analytical process. Finally, section 3.6 discusses the problems found during the analysis and puts forward the solutions that have been implemented to overcome such difficulties.

Chapter 4 displays the analysis itself and introduces the findings and data obtained after applying the methodologyical steps described in chapter 3. As it has been designed, the stepwise methodology of this research has to go through the following stages. First, the definition of search codes; second, the implementation of searches; third, the definition of filters and finally the refinement of the results. This ordering, as well as the fact that the three parts of the word, namely derivational prefix, stem and inflectional ending, are subject to variation and require specific analysis, laid the foundation for the organisation of the chapter. Thus, Section 4.2 concentrates on the search commands and the results obtained. Subsection 4.2.1 provides a summary of the search commands whereas subsections 4.2.2 through 4.2.8.8 display the unlemmatised forms found for each verb class and subclass with each of the following search command and search filters. Section 4.3 shows the results of the analysis of the elements at both sides of the stem. Thus, 4.3.1 depicts the situation of preverbal items while 4.3.2 concentrates on the inflectional endings. Finally, subsection 4.3.3 constitutes the goal of the research and displays the new list of lemmas obtained from the analysis. Finally, some considerations regarding the choice of lemmas are made in 4.4.

Chapter 5 offers an overview of the tasks that have been undertaken for this doctoral dissertation and offers potential lines of research to be followed in the future. Thus, section 5.2 offers a summary of the previous chapters. Section 5.3

discusses the main conclusions of this work at qualitative (5.3.1) and quantitative (5.3.2) levels. Finally, 5.4 displays the goals achieved by this thesis and discusses future research venues.

2. THEORETICAL BACKGROUND

2.1. Introduction

Chapter 2 sets the lexicological basis of this work. As has been explained in the introduction, the aim of this doctoral dissertation is to offer a lemmatised inventory of the strong verbs of Old English based on the *Dictionary of Old English Corpus*. Such a task, which falls on the side of lexicography rather than on lexicology only, requires a solid lexicological foundation, to which this chapter is devoted. In other words, this chapter provides the theoretical aspects of the research while the next one insists on the applied side of the question.

This chapter opens with some general remarks on the linguistic field of lexicology, including both the non-derived and the derived lexicon, (section 2.2) and goes on to discuss some specific aspects of the lexicology of Old English (section 2.3). Special attention is paid to the verbal system and more particularly to strong verbs (section 2.4). In so doing, the origins and evolution of the strong verb system are analysed, from Indo-European (Section 2.4.1) through Germanic (section 2.4.2) to Old English (section 2.4.3). This section depicts the general structure of the verbal system in Old English including weak, strong and anomalous verbs. In the following subsections, the Old English strong verb system is fully described. First, section 2.4.3.1 discusses the traditional classification of the strong verbs, along with some alternative classifications. The inflectional paradigm of the strong verbs including dialectal variation is depicted in section 2.4.3.2, whereas derivation originated in the strong verb is presented in section 2.4.3.3. To close this chapter, some concluding remarks are made in section 2.5.

2.2. An overview of lexicology

Lexicology is a branch of linguistics and, as such, it essentially adopts a theoretical perspective on the study of words: their form, function and meaning along with the processes by which new words are coined. This area of linguistics is closely connected to semantics, morphology, syntax and pragmatics. Semantics is the branch of linguistics in charge of the study of encyclopaedic meaning in language. Syntax analyses the functions and position of the elements of a language in an utterance. Morphology focuses on the study of word formation. These two branches are closely linked, especially in inflectional languages, where inflections indicate the functions the elements have in a particular utterance or sentence. Pragmatics concentrates on the real use of the language, especially as regards implied meaning. It is precisely the meaning which is not derived directly from the real meaning of the words used which becomes relevant for lexicology, and particularly to lexicography, as a dictionary needs not only to include encyclopaedic meaning but also the meanings that the speaking community may have added to certain linguistic elements to cover some concepts that may not be canonical.

In semantics, the process of lexicalisation plays an important role, as this process extends the meaning of words through word formation, the creation of words to account for a new concept, or through the extension of the meaning of an existing word to cover this new concept. In Brinton and Traugott's (2005: 32) words, lexicalisation is "the process by which new items that are considered *lexical* (in terms of the theory in question) come into being".

Brinton and Traugott (2005) focus on the changes languages face throughout their existence, and the adaptations that words suffer to get the meaning that they currently convey, as well as the new uses words acquire as the speaking community evolves. Along with the description of the changes, Brinton and Traugott (2005) provide the reader with a theoretical account of the concepts that explain such changes.

Thus, Brinton and Traugott (2005) provide a general definition of the main lexical creation processes they study in their work. According to Brinton and Traugott (2005: 32), there are different processes of lexicalisation, including:

1. Ordinary processes of word formation.
2. Processes of fusion resulting in a decrease in compositionality.
3. Processes of separation resulting in an increase in autonomy.

The processes Brinton and Traugott (2005: 34-44) include as general translinguistic lexical creation tools are:

Compounding: Brinton and Traugott (2005: 34) follow Bauer (1985) and describe compounding as a process that "involves the unifying of two or more autonomous words to form a third". After this process takes place, the parts forming the new-resulting word are no longer independent, as shown by phonological changes (shift of stress) or the loss of semantic motivation, as shown by *blackboard* (N) < *black* (Adj) + *board* (N), where the stress is diminished in the second syllable and the colour of the product is not restricted to black. Brinton and Traugott (2005: 34) also point out that in inflectional languages, as it is the case with Old English, "it is preferable to think of the combinations between *roots* or *stems*, rather than words (where *root* is a form that carries primary lexical meaning and cannot be analysed further, and *stem* is a lexical form minus inflectional morphology)". The existence of synchronic rules for compounding may lead to the loss of the understanding of the relationship between the complex word and its constituents, thus bringing opacity upon the original productivity.

This may be well represented by the unmotivated phonetic realisations of compounds like *cupboard* /ˈkʌb.əd/ or *forecastle* /ˈfɔrˌksəl/.

Derivation: it results in the formation of new lexemes, by adding a derivational affix to a root or stem. However different they may be from a theoretical point of view, the distinction between compounding and derivation is not always

clear-cut. Rather, fuzzy areas emerge, especially when derivational morphemes are given lexeme status (as happens to be the case with *-ism* and *-ology* in Present Day English, or when lexical items engage in an unstoppable process of grammaticalisation as was the case of *dōm* or *hād* in Old English. Derivational morphemes are traditionally grouped in two sets, namely the *purely semantic/lexical* affixes which add lexical meaning to the root they are attached and the *grammatical* affixes, which only incorporate grammatical meaning. In the first group belongs *un-* (meaning 'not' in *unnecessary* or 'reversal' in *undo*) whereas the second group holds affixes such as *-er* which implies an agentive function in words like *singer*.

All the above-mentioned cases fall within the scope of derivational morphology. However, bound morphemes may also be inflectional, and not only derivational. Brinton and Traugott (2005: 35-36) state that even if the exact distinctions cannot be clearly established, there is at least general consensus on the fact that derivational affixes differ from inflectional affixes in the following features:

1. Grammatical derivational morphemes are not obligatory, i.e.: agentives in English can be marked by *-er* or not (*swimmer vs. chairperson*)

    1. They compete with other derivatives with the same function. Nominalisers like *-ness* and *-ity*. E.g.: *pettiness*; *rigidity*, but **pettity*, **rigidness*.
    2. They may involve variables in a relatively open system. Affixes of different types may be attached to the same root to produce different results (nominalisations, adjectivalisations or other verbalisations).
    3. It can end up in a syntactic recategorisation, and the creation of new bases of derivation.

By contrast, inflections typically:

4. Are obligatory. i.e.: In English, inflection is needed to mark the verb-subject agreement (she walks, they walk).
5. Compete only marginally with other inflections having the same function (the past participle –en of strong verbs vs. the past participle d/t/-ed of weak verbs seen/ridden vs. slept/dreamed). This competition is well represented by originally strong verbs that have had there past participle assimilated to the weak paradigm as is the case with dive/dove/dived *doven).
6. Involve variables in a relatively tight, closed system (singular and plural forms).
7. Do not change class membership or produce new lexemes.

However, some inflections are problematic and exemplify the continuum derivation-inflection. According to Haspelmath (2002: 230) present participles may behave as adjectives with respect to word order and syntactic modification of nouns, but also as verbs as they may combine with different complements. In the specific area of Old English studies, González Torres (2010a, 2010b, 2011) has identified some areas of overlapping between inflection and derivation that affect the morphemes *-a*, *-e*, *-o* and *-u*.

Conversion: Following Bauer (1983) Brinton and Traugott (2005: 37) state that conversion "is a *functional shift* from one category to another equated in English with zero derivation". This process typically involves derivation from one major lexical class item to another. Traditional examples of conversion include run (N) < to run (Vb); private (N) < private (Adj); to lower (N) < lower (Adj). In the case of English, it is usually equalled to zero derivation, but as it is the case with languages with little inflection, we may see examples of conversion from one minor lexical class into a major one, as in to off (Vb) < off (Adv); if (N) < if (Conj) in the expression *ifs, buts and ands.*

In languages with inflection, conversion can be more derivationally complex as seen in the cases of the German verb *duzen* < *du* (Pron) or the French verb *tutoyer* < *tu* (Pron). The case with prepositions has been treated

separately and, according to Zielgeler (2003: 229), they are "better seen as a type of metonymy, motivated by a situation in which the particle or prepositional form […] can in time serve to stand for the entire verb phrase", as it is the case in the Singapore English sentence *to off the microwave*. A shift from a minor class to a major class is widely recognised as lexicalisation (and degrammaticalisation; Norde 2009). In this respect, Hopper and Traugott (1993: 49) define lexicalisation as "the development of a fully referential lexical item from a non-lexical, or grammatical item, such as the development of the verbs *up, down* or nouns *upper*, *downer* from the homophonous particles *up, down*".

Clipping and ellipsis: Both processes are understood as means for lexicalisation. In both cases there is a loss of lexical material. According to Blank (2001: 1605-1606), they differ in that "[clipping is the result of] - DMR the deletion of one or more syllables from multisyllabic words, whereas ellipsis leads to the formal reduction of a complex word or phrase". In the case of ellipsis semantics plays an important role as "the semantics of the omitted element is absorbed into the remainder by metonymy" Brinton and Traugott (2005: 40). Examples of clipped words include *flu* < *influenza*; *fan* < *fanatic* or *fridge* < *refrigerator* whereas *narc* < *narcotic agent* or *pub* < *public house* account for cases of ellipsis

Blending: it "involves the fusing of words into a single lexeme, or *portmanteau word*, by a process of compounding and clipping" (Brinton and Traugott 2005: 41). Blending may cause the destruction of the integrity of both elements, or may mix together the beginning or end of one element leaving either the first or second element intact.

Instances of integrity destruction are *bit* < *b*(*inary*) (*dig*)*it* and *heliport* < *heli*(*copter*) (*air*)*port*. The first element is preserved in the words *filmography* < *film* (*bi*)*ography* and *skyjacker* < *sky* (*hi*)*jacker*; whereas the second element remains intact in the cases of docudrama < *docu*(*mentary*) *drama* or *blog* (*we*)*b log*.

Back formation: it is, in the words of Brinton and Traugott (2005: 41), "the creation of a morphologically simple form from a word which is analysed (frequently incorrectly) as a morphologically complex word on the basis of analogy with derivational and inflectional patterns existing in the language".

Back formations based on interpreted derivational forms include the verbs *enthuse* < *enthusiasm* and *laze* < *lazy* whereas back formations resulting from the misinterpretation of a simplex word as an inflected form include pea < Fr. *pease* (sg.) or skate < Du. *Schaats* (sg.).

Acronym: it is a lexicalisation process by which "the initial sounds (or letters) of words within a complex word or phrase are put together into a unified lexeme" (Brinton and Traugott 2005: 42). They may be represented in lower case, thus appearing as fully formed morphemes or in upper case and thus not so easily interpreted as lexemes. Examples of acronyms used a proper lexemes are *laser* < *l(ight) a(mplification (by) s(timulated) e(mission) (of) r(adiation)* and *radar* < *ra(dio) d(etecting) a(nd) r(anging)*, whereas *AIDS* < *a(uto) i(mmune) d(efficiency) s(yndrome)* and *NATO N(orth) A(tlantic)T(reaty) O(rganisation)* stand up as cases of less obvious lexemes.

Closely connected to the notion of acronysm is that of *initialism*, which involves the articulation of the initial letters of words within a complex word or phrase as letter forming a new word. Initialisms are to be represented by IBM < I(nternational) B(usiness) M(achiness) or OED < O(xford) E(nglish) D(ictionary).

Writing conventions, either upper case or lower case with periods, suggest that initialisms are not fully formed words (Briton and Traugott 2005: 42), their choice may imply a semantic difference as in AM < amplitude modulation *vs. a.m.* < *ante meridiem.*

Loan translation or *calquing*: Calques are "literal translations of terms from another language. In the source language the terms may themselves be derived or partial borrowings, and once they are calqued, may undergo normal

processes as phonological reduction" (Brinton and Traugott 2005: 43). Examples of borrowing in the history of English include OE *fore-set-nyss* 'in-front-set-ness' < Lat. *praepostio* 'preposition' and OE *betwux-aleged-nys* 'between-put down-ness' < Lat. *interjetio* 'interjection'. In these cases, the Latin terms have eventually substituted the calques, but some Old English calques have survived and have undergone phonological and semantic changes like *gospel* < OE *gospell* < *god* + *spell* < Gr. *evangelion* 'good tidings' or *Holy Ghost* < OE *Halig Gast* < Lat. *Spiritus Sanctus.*

Coinage or *root creation*: it is an exceptional process that "involves the creation of a new root morpheme" (Brinton and Traugott 2005: 43). Coinages may be *arbitrary* or *motivated*. Bussman (1996) considered coining to be fully arbitrary while McArthur (1992) states that coinages may be *motivated* or created *ex-nihilo*. Motivated coinages include onomatopoeic words, such as *gulp* or *hiss* while typical arbitrary coinages, those in which "there is no lexicological way of accounting for the formation of a word" (McArthur 1992: s.v. "root-creation") are represented by the words *Kodak* or *googol*.

Metalinguistic citation: it refers to "the speaker's ability to pick up any piece of linguistic material and make it into a word" (Brinton and Traugott 2005: 44). Heine (2003b: 166) names this process *revaloratiation* and states that it "serves to assign a segment of text, irrespective of whether it is a meaningful entity or not, the status of a referential, lexical form". A prototypical example is *there are two* e'*s in my name*, where the letter *e* has is used in the plural, and thus promoted to the category of noun.

This lexicological review of word formation processes provides a list of tools that clearly surpasses the resources of which Old English speakers made use at the time. In this work, only the processes of compounding, derivation, zero derivation and inflection will be taken into consideration. However, before entering the description of the inflectional and derivational paradigms of the strong verbs, a word of caution must be given regarding the general features of Old English and the reconstruction of lexical elements in a dead

language. The following section will focus on the importance of the morphophonemics of Old English as a basis for the classification of its different paradigms and lexical reconstruction.

2.3. The morphophonemics of Old English: A path for lexical reconstruction

Old English is the name given to a stage of the language that comprises more than five centuries. In this respect, it is not a uniform and well-set language. Rather, it is very much subject to change in the different areas that make up a language. Because of historical inheritance, internal and external factors, the expansion of the language across regions and the contact and influence with other languages, the panorama of Old English lexicography is far from being easy to cope with. This section aims at pointing out some properties that must be borne in mind when dealing with Old English and trying to synthetise its inflectional systems and normalise its spelling. In general, works on Old English focus on specific areas of the language rather than attempting to offer a comprehensive view of the language. In this sense, perhaps Hogg (1992) constitutes the most comprehensive work on the Old English language up to the present. Hogg (1992) gives an exhaustive account of the history of English, covering all the fundamental aspects, from orthography to syntax, including grammatical categories and semantics. On the grammatical side Hogg (1992) assesses the lack of reliable grammars from Old English times because "the Anglo-Saxons would not have wished to make such a claim, their intellectual interests lying in entirely different areas" (Hogg 1992: 67). The only proof of grammatical writing in that period is Ælfric's *Latin Grammar*. For his purpose of study, there is no reference work to reconstruct the nature of Old English. For so doing, one must rely on the remaining textual material and check it against the different branches of linguistics. For the purpose of this research I will concentrate of the features that are most clearly related to morphology, namely orthography and phonology.

On the orthographical side, the main important feature Old English is the use of the Roman alphabet, together with the runic alphabet, which the Anglo-Saxons brought with them. The runic alphabet was only used for inscriptions and dedicatory formulae rather than for communication. In Old English times, conventions for orthography varied from scribe to scribe, so what we have nowadays was unknown in that time. The Roman alphabet was adopted by the Kentish king Æthelbert as soon as the first manuscripts written with the Roman alphabet appeared. This showed the spread of Christianity in that time. Writing was a property of the church, and instruction in reading and writing was a monastery duty, and normally scribes were clerics. Even the letter style depended on the monasteries.

The Roman alphabet used for writing Old English was a variant of the general Medieval Latin one, augmented by five characters of runic origin:

(1)
a. *þ* thorn
b. *ƿ* wynn
c. *ð* eth
d. *æ* ash
e. *ȝ* yogh

The Roman alphabet had many advantages compared with the runes: it was common to all western Christendom; and it was already adapted, as runes were not, for use with quill and parchment. Its rapid adoption was in no way surprising. Minuscule script, for example, was introduced in the early 8th Century in the Latin text *Lindisfarne Gospels*. Yet the runic alphabet was mainly used for inscriptions, and was used until the 11th Century. It was capable of continual modification and adjustment, which means that it was very flexible. The changes in the runic alphabet are characterised by: changes coming from the nature of their interaction with one another. If there was a change in pronunciation in the language, the rune came to represent the new sound. This means that there was an understanding of sound changes in the

development of the language. Yet, phonological reconstruction is not straightforward, and we can only rely on partial evidences. In Hogg's (1992: 69) words, "we can assume that the spelling conventions in use during the Old English period would not have been completely at odds with those in later periods". There is not enough spelling evidence of the Old English texts, "given the Old English word *bedd* it is reasonable to assume that the first letter represents some kind of consonant, the second some kind of vowel, the third some consonant different from the first. On the other hand the fourth letter causes us problems, since it could well represent a departure from present-day orthographical conventions" (Hogg 1992: 68). On the contrary, dialectal written evidence presents us with a greater range of information than that of the standard language, their forms being derived directly from the standard Old English language. In addition to this, the dialects define limits of reconstruction, given the fact that "one general principle of linguistic reconstruction is that we should not hypothesise for older stages of the language forms which do not appear in later stages" (Hogg 1992: 69). Another type of evidence comes from *linguistic plausibility*, the probability that exists, if we reconstruct a linguistic system for one stage of the language, to account for the differences between this stage and a later one by possible linguistic changes. The last source of evidence comes from the relations between Old English and contemporary Germanic languages in the continent, and the borrowings Anglo-Saxons took from other languages, especially Latin. It is important to note that, by the analysis of these borrowings from Latin, we can reconstruct the phonologic conventions in Old English, as they borrowed words from Latin maintaining their form, and most probably, their phonology and pronunciation. It is important to note that none of these types of evidence can exclude any of the others. The process of reconstructing a linguistic system is uncertain and all sources of information contribute to the final aim, which should be that of knowing the distribution and interaction of the elements in the language, and more concretely, the reconstruction of phonemes or contrastive sounds of the language.

On the morphological side, reconstruction poses questions different from those of phonology, although some problems are still related to phonology, as is the case with *i*-mutation, illustrated in the following example: "Consider PDE *foot* - *feet* […] we can for proto-Old English reconstruct singular */fo:t/, plural */fo:tiz/. The /i/ in the second syllable of the plural then caused the vowel of the first syllable to become a front vowel, eventually emerging as /e:/, so that we ding singular *fōt*, plural *fēt*, and hence the present-day forms" (Hogg 1992: 71). Another problem associated with morphology would be that of cataloguing verbs different morphologically into the same grammatical class. Take the following example as an illustration "there can be no doubt that OE *swimman* 'swim' […] and *drincan* 'drink' belonged to the same morphological group or class of verbs […] But in Old English we have other verbs such as *helpan* 'help', *weorþan* 'become', and *berstan* 'burst' which, although their patterns are similar to that of *swimman*, nevertheless show clear differences in the quality of their stem vowel" (Hogg 1992: 71). In most occasions phonological and morphological problems overlap.

Some of the properties of Old English are direct consequences of phonological and morphological changes that took place in Proto-Indo-European or in early Germanic and are sometimes difficult to reconstruct. Let us concentrate on the verbal paradigms. The Germanic languages developed a dual system of verbs: strong and weak. The difference lies on the formation of tenses: strong verbs form their preterit by means of vowel variation, while weak verbs form their preterit by the addition of a dental suffix. There is also room for irregular verbs, both preterit present verbs and anomalous verbs. The strong conjugation is the oldest, having its origin in Indo-European, while the weak conjugation was originated in the Germanic languages and became productive. All the new verbs in Old English are inflected following the weak conjugation paradigm.

The following section discusses the strong verb system in Indo-European and Germanic and provides a general overview of the whole verb system in Old English with special attention to the classification and characteristics of strong verbs.

2.4. The strong verb system: A historical overview

This section engages in the description of the historical evolution of the strong verb system, from the language in which it originally developed, Indo-European (2.4.1), through Germanic (2.4.2) to Old English (2.4.3). Throughout time, the productive, strongly systematised and phonologically conditioned system of Indo-European was slowly dismantled. In Germanic times, a second verbal system was created that became productive during the Old English period. Yet, the strong verb system survived and some of its features were preserved. In the following subsections, the process of collapse of the system and the reasons for that outcome will be revised. Since Old English is the subject of study, special attention will be paid to the coexistence of both the weak and strong systems, the classification of the strong verbs and their inflectional and derivational properties (sections 2.4.3.1, 2.4.3.2 and 2.4.3.3).

2.4.1. The origins of the strong verb system: A Proto-Indo-European inheritance

In order to stress the relevance of the strong verbs for the organisation of the Old English lexicon, in the following a comprehensive view is offered of the origins and evolution of this verbal system through the history of the different related languages. Mailhammer (2008) concentrates on the importance of the *ablaut* patterns which rule the strong verb paradigm since Proto-Indo-European. *Ablaut* is a morphological process that alters the root vowel in a way to mark a different grammatical or inflectional role in the language. According to Birkhan (1985:32) *ablaut* "is defined as the regular alternation of vowels in roots and affixes of words that are etymologically and morphologically related". Mailhammer (2008: 187) considers that "ablaut as a phenomenon cannot only be found in the Indo-European languages but also in other language families, most notably in Semitic, which possesses an extensive and functional ablaut system". So, in Mailhammer's (2008: 186) words, two

questions come to mind, “what exactly is the Germanic innovation, and how can it be adequately modelled? Second, what is its typological and diachronic significance?” From his point of view, the ablaut pattern in the Germanic strong verbs comes from Proto-Indo-European, but is more widely and more successfully adapted in Proto-Germanic.

When talking about the origins of ablaut in Proto-Indo-European, three questions arise, “is the change phonologically relevant, and does it therefore result in allomorphy? […] Does the change constitute a distinctive feature, or is it used redundantly? […] Is the change conditioned phonologically or morphologically?” (Mailhammer 2008: 189-90). Sometimes in Proto-Indo-European, ablaut changes do not necessarily reflect grammatical and semantic relations. Take Mailhammer’s (2008: 190) example as illustration:

> “the verbal roots of PIE +mad- ‘become’ and +med- ‘measure’; they constitute two lexical entities that are neither etymologically nor morphologically related. This means that the different root vowel does not indicate a semantic or grammatical relationship between these two roots”.

Apparently, the radical vowel in Proto-Indo-European “is important for the lexical identification of a root” (Mailhammer 2008: 191). The conclusion that can be drawn is that ablaut in Proto-Indo-European did not depend exclusively on phonology, but rather on morphology.

In understanding how verbal paradigms are constructed, Mailhammer (2008:191-192) states that:

> “A verbal root by virtue of its semantic denotation (its lexically determined meaning) automatically belongs to one of two aspectual poles. The imperfective aspect (traditionally called *present*) refers to an atelic mode of action (*aktionsart*), whereas the perfective aspect (traditionally called *aorist*) denotes a telic mode of action”

This means that there are two types of verbal roots in Proto-Indo-European. The problem is that verbs do not express only one aspect (present or perfective), but there are different types of stem formation, which is defined as

a “pre-defined template for the construction of a paradigm, determining the morphological properties that are used” (Mailhammer 2008: 192). There are two kinds of stem formations: thematic and athematic. Athematic ones do not add a suffix to the verbal root, instead they add a personal ending directly after the root. Athematic stem formations display two different stems: strong and weak, with different radical ablaut grades and different patterns of accentuation. Thematic formations add a characteristic suffix to the root and do not have radical ablaut. Rather, the vowel of the suffix depends directly on apophonic change. In conclusion:

> “a verbal root in PIE generally has at least two separate aspectual paradigms, each formed according to the rules of a different type stem formation. In addition to this, some roots also form a third paradigm, the perfect, which belongs to the imperfective aspect […] the perfect differs from the present and the aorist in that for each root there seems to exist only one possible type of perfect stem formation per type of root” (Mailhammer 2008: 193).

If we take into consideration the hypothesis that ablaut can be considered a distinctive morphological property, then “the expression of grammatical categories is based on verbal stems that reflect the dimension of aspect […] a verbal root usually forms a present stem and an aorist stem. In addition, some roots possess a perfect stem as well as secondary formations, such as intensive stems” (Mailhammer 2008: 193). Stems are formed following complex morphological rules, namely ablaut, affixation, reduplication and word stress. The combination of the morphological properties in a verbal stem specifies the type of stem formation. Ablaut on itself cannot allow the identification of a type of stem formation, it is a redundant property associated with a given type of stem formation. If this process cannot differentiate between different types of stem formation, it cannot express the category of aspect in the verbal system in Proto-Indo-European.

2.4.2. The Germanic strong verb system

Ablaut is directly related to the Germanic languages, and, more specifically, to the strong verbs in Germanic languages. Ablaut is systematic in the sense that it creates a high degree of organisation in the strong verbs. Thanks to this, strong verbs constitute a strong system, and ablaut in this system is the basis of strong verb inflection. In this respect, there is still debate on how it works as compared to the parent language (Proto-Indo-European), and on how both languages differ. According to Mailhammer (2008: 199), there are two opinions: some scholars accept "the importance of ablaut for the Germanic verbal system, but they do not see a typological difference to the parent language" while some other authors state that there is "a unique extension of ablaut in Germanic, without going any further. […] This general notion remains obscure, because the exact status of ablaut is not sufficiently conceptualised […] it is unclear in what way the systematisation and functionalization of ablaut in Germanic affects its typological position".

When compared with Proto-Indo-European, in which grammatical categories are expressed by complex morphological rules, the system of strong verbs in Germanic works on the basis of a verb's root structure, with the exception of the preterit presents, not always identified by their root structure. The corner stone on which strong verbs in Germanic are organised is a verb's root structure, "which determines its stem formation via the predefined paradigmatic change of ablaut grades" (Mailhammer 2008: 201). A given Germanic strong verb has up to four different stems, differentiated by different ablaut grades. It follows that "the dominant position of ablaut in the strong verbs is illustrated by the fact that the whole system was organised around this morphological property" (Mailhammer 2008: 201). One of the main characteristics ablaut has in the Germanic strong verbs is that it allows for the classification of verbs depending on the root vowel resulting from ablaut, that is, "the strong verbs are subdivided into a *Primary System* (traditional class I to V or van Coetsem's *e*-verbs), a *Secondary System* (class VI or a large part of van Coetsem's *a*-verbs) and a parallel system of verbs acting as a 'safety net'

for verbs that could be integrated in neither group" (Mailhammer 2008: 202). Thus, the first two systems comprise ablauting verbs, and the parallel one includes reduplicating verbs. The reduplicating verbs appearing in the parallel system stated by Mailhammer (2006) cause the apparition of the Class VII in Old English, "this new class is characterised by an ablaut-like alternation of the root vowel that displays one ablaut grade for the stems of the present and the past participle and another one for the stem of the preterit" (Mailhammer 2008: 205). Ablaut creates subdivision in the systems, as it continues from Proto-Germanic because of regular sound change, due to this, Old English Class III is subdivided into three subclasses, depending on the consonant that follows the root.

As the inflectional system collapses, the number of forms distinguished by means of ablaut alone increases. This phenomenon becomes more relevant by the evolution of class VII, and ablaut becomes more functional. However, the whole system of strong verbs becomes, at the same time, fragmented by the regularisation of sound change, which will end up in the formation of a weak inflectional pattern. Consequently, "the trend is towards an increasing functionality of ablaut, while the system as a whole is weakened by fragmentation" (Mailhammer 2008: 206). Ablaut in Proto-Germanic is systematised and functionalised, and in stem formation ablaut becomes a functionally distinctive aspect, which can express grammatical categories.

Mailhammer (2007) revisits the notions of ablaut and reduplication by specifically focusing on the Germanic strong verbs. He distinguishes two kinds of ablaut: quantitative and qualitative. Mailhamer (2007: 16) identifies quantitative ablaut in those cases where "the ablaut vowel appear either unaltered (full grade), lengthened (lengthened grade), or not at all (zero grade)". As regards qualitative ablaut, he supports the idea that it developed through accent. In Mailhammer's (2007: 17) words, "the musical accent of Indo-European was the reason for the *Abtönung* of full grade and lengthened grade vowels" (Mailhammer 2007: 17). The second feature that characterised the Germanic strong verb system is reduplication. Sapir (1921: 126) defined reduplication as "the repetition of all or part of the radical element". Marantz

(2000: 557) proposes a more wide definition and states that reduplication is “a type of word formation […] in which the phonological form of an affix is determined in whole or in part by the phonological form of the base”.

These two processes take place in Germanic strong verbs, but ablaut is more generalised. Reduplication only occurs in the preterit of the reduplicating verbs, but the majority of the strong verbs create the preterit tense through ablaut, something supported by the idea that “since the preterit of the ablauting verbs (classes I to VI) in the Germanic daughter languages show no trace of reduplication, it therefore has to be assumed that they were already unreduplicated in the Germanic parent language” (Mailhammer 2007: 34). This double pattern, leads to the subclassification of strong verbs into ablauting and reduplicating verbs. The second group of verbs are further subdivided into verbs that perform both ablaut and reduplication, and those that only use reduplication. Reduplication has been assumed to take place only in perfect formation in the parent language, namely in the Indo-European perfect forms.

Contrasted with reduplication, ablaut takes place more commonly in the Germanic strong verbs. Its position was strengthened when it became systematised and functionalised. As Indo-European does, Germanic combines quantitative and qualitative ablaut in one paradigm. Although this is generally accepted, there are two strong opposite opinions regarding the degree of organisation the strong verbs attained through the systematisation of ablaut, “there are authors who acknowledge the importance of ablaut for the Germanic verbal system, but they do not see a typological difference to the parent language” (Mailhammer 2007: 46). Some other authors “notice a unique extension of ablaut in Germanic, albeit without going any further […] it is unclear in what way the systematization and functionalization of ablaut in Germanic affects its typological position” (Mailhammer 2007: 46).

As stated above, Mailhammer (2007, 2008) acknowledges the existence of a primary and secondary classification of the Germanic strong verbs. To be precise, the primary system gathers all the Germanic verbs that presented an *e* as their root vowel. At early stages of this period, all verbs having an *a* or an *e*

as root vowels were grouped together in this primary system, as they presented the same ablaut pattern "*e*-grade for the present (indicative, subjunctive, participle and imperative), *a*-grade for the preterit singular and zero grade for the remaining form (preterit plural, preterit subjunctive and past participle). It is unclear, though, that this classification was so, because it is not known if all *e*-verbs had *e* as a root vowel over the entire period of Germanic. This primary system was eventually divided into five different classes, the reason being that "early sound changes caused a split into several different subgroups, the well-known ablaut classes" (Mailhammer 2007: 58).

Thus, Class I is characterised by a root structure *consonant+e+i+consonant*, a diphthongal base in which the main importance falls on the radical *e*, as it is upon it that the ablaut operates. The characteristic diphthongal root structure automatically determines its inflection according to the pattern of Class I. This class shows the regular ablaut pattern of the primary system, but it has one irregularity: at some stage of Germanic the root vowel becomes monophthongal. Morphologically and semantically speaking, there is one feature present in class I strong verbs, that being that it contains verbs possessing similar root structures as well as similar meanings as shown by **sleika-*, **sleida-*, **gleida-* and **skreipa-* all of them meaning 'to glide'. Seebold (1970) provides a full account of this phenomenon.

Class II has a very similar root structure to class I, although the difference is in the second element of the diphthong (*i* in class I, *u* in class II), thus the pattern is *consonant+e+u+consonant*. The ablaut pattern is the same, although in the fourth stem (past participle) -u becomes -o as a result of breaking in the attested daughter languages.

Class III, which presents as the characteristic root structure *consonant+e+sonorant+consonant* has a special position in the Germanic strong verb system. The author remarks "the syncretism leading to the development of the system of strong verbs" (2007: 63), and follows Prokosch (1939) on stating the heterogenic origin of the verbs in this group. According to Prokosch (1939: 169) "every verb of this class may be considered a transfer from some other class, either through extension by a determinant or through a

nasal infix". It is only due to homogenisation that class III shows a uniform ablaut pattern.

Class IV has as a characteristic root structure *consonant+e+sonorant* presents a similar ablaut pattern as the previous strong verb classes. Only the preterit presents show the regular zero-grade in the present plural stem, as the rest of the verbs took the lengthened grade from class V. In contrast, some verbs that present a root structure similar to class V have a similar pattern with class IV in some dialects.

Class V shows a canonical root structure, *consonant+e+consonant*, but this causes some changes in the regular ablaut pattern. A number of verbs that have been seen as class V verbs are actually grossed with class III in the attested Germanic daughter languages. These verbs are those with two root-final obstruents.

Contrasted with the primary system, the secondary system is characterised by a different root vowel, a different ablaut pattern and a different chronological development. Compared with the primary system, the secondary system shows only two different ablaut grades, which give way to two different stem forms: present tense, past tense; and preterit. The author states "the traditional representation of class VI merely as a variant of the other five classes is inadequate" (Mailhammer 2007: 87). There are two features that differentiate the primary system from the secondary one.

First, the basic root vowel is always *a*, followed by a single consonant or a cluster, in which the first element is an obstruent. Second, the ablaut pattern is different from that of the primary system, much simpler due to "the absence of the characteristic and inherited difference in ablaut grade between preterit singular and plural" (Mailhammer 2007: 88).

Finally, the reduplicating verbs constitute a distinct group among the Germanic strong verbs. They are "preterits with partial reduplication of the root syllable, and can be unambiguously identified via the structure of the present tense stem, i.e. the root" (Mailhammer 2007: 104). Ablaut was not possible in these verbs as they presented "heavy bases with either a diphthong or *ō* as their root vowel" (Mailhammer 2007: 104). The formation of an ablaut

grade was impossible due to phonotactic reasons. Instead, reduplication was used for the characterisation of the temporal opposition. Some authors claim that "reduplication is a Germanic innovation based on the *verba pura* with long vowels, as opposed to other authors who have posited that reduplication is an archaic feature" (Mailhammer 2007: 105).

All in all, strong verbs in Germanic constitute a highly organised class as a result of the systematisation of ablaut. Compared with the parent language, verbs in Germanic are classified on the basis of their phonological root structure, stem formation being negligible from a synchronic point of view. Secondly, ablaut in the system of strong verbs was functionalised in a unique way if compared to other Indo-European languages. The root vowel of a strong verb has a functional value, determining the assignment of the relevant subsystem and the ablaut class.

2.4.3. The Old English verb system

This section offers an exhaustive description of the morphophonological features and structure of the verbal system in Old English, with special attention to the strong verbs. Weak verbs and anomalous verbs are also discussed. It must be borne in mind that the description that follows in the remainder of this work, unless indicated, makes use of graphemes, even though angles are avoided for simplification.

Old English inherited much of the verb structure present in Germanic. However, the passing of time and interrelated processes of weakening and levelling gave way to the collapse of the strong verb system and the growth and standardisation of the weak verb system. In fact, the weak system became the only productive system by the end of the period.

The Old English verb had a much simpler conjugation than other Indo-European languages such as Greek or Latin. Morphologically, the Old English verb showed distinctions of mood, time, person, number and voice. There were four moods: indicative, subjunctive, imperative, and infinitive; the number

distinctions involved the singular and plural, having lost any trace of the dual; as regards time only two simple tenses were distinguished by means of inflection: present and past. Three persons were recognised, although some of the endings were coincidental; with respect to voice, there were an active and a passive voice although the passive had not a morphological realisation and was generally expressed by means of the auxiliaries *wesan* 'to be' and *weorþan* 'to become' followed by the past participle**.** However there is a trace of morphological passive in the forms *hatte* 'is/was called', and *hatton* 'are/were called', but this seems to be an analogical formation with the singular *hatte*.

To illustrate the inflection of a weak verb, Figure 1 displays the paradigm of *cepan* 'to observe, keep, regard':

	Infinitive	**Preterit**	**Past Participle**
	cepan	cepte	geceped
	Present	**Preterit**	
1st sg.	cepe	cepte	
2nd sg.	cepest	ceptest	
3rd sg.	cepeð	cepte	
Pl.	cepað	cepton	
	Subjunctive	**Subjunctive**	
Sg.	cepe	cepte	
Pl.	cepen	cepten	
	Imperative		
Sg.	cep		
Pl.	cep		
	Inflected Infinitive	**Present Participle**	
	cepenne	cepende	

Figure 1: Inflectional paradigm of the weak verb *cepan* 'to observe, keep, regard'.

The prototypical paradigm of strong verbs is illustrated in Figure 2 by means of the inflection of the verb *dreopan* 'to drop':

	Infinitive	**Pret. Sing.**	**Pret. Pl.**	**Past. Part**
	dreopan	dreap	drupon	dropen
	Present	**Preterit**		
1 sg.	dreope	dreap		

2 sg.	driepst	drupe
3 sg.	driepþ	dreap
Pl.	dreopaþ	drupon
	Subjunctive	**Subjunctive**
Sg.	dreope	drupe
Pl.	dreopen	drupen
	Imperative	
Sg.	dreop	
Pl.	dreopaþ	
	Inflected Infinitive	**Present Participle**
	dreopenne	dreopende

Figure 2: Inflectional paradigm of the strong verb *dreopan* 'to drop'.

As can be seen in figures 1 and 2, the Old English verbs, as occurs in all Germanic languages, are divided into two groups, depending on the formation of their preterit:

8. Weak verbs: which form their preterit by the addition of the dental suffixes *-da, -ta*.
9. Strong verbs: which form the preterit form by means of ablaut.

The weak verbs are classified around three citation forms, namely the infinitive, the preterit and the past participle, and are grouped in three main classes. Class I is characterised by the *-de/-te* ending in the preterit and and *-ed/-t* in the participle. Class II is recognisable by the *-ian -ode* and *-od* endings of the infinitive, the preterit and the participle respectively, whereas class III includes only four verbs of a special character, those being *habban* 'to have', *secgan* 'to say', *hycgan* 'to think' and *libban* 'to live'. A sample of the weak verb classes is presented in Figure 3:

	Infinitive	**Preterit**	**Past Participle**
I	fremman 'to do'	fremede	gefremed
	cepan 'to keep'	cepte	geceped
	hieran 'to hear'	hierde	gehiered

	ferian ‘to carry’	ferede	gefered
	bycgan ‘to buy’	bohte	geboht
	þencan ‘to think’	þohte	geþoht
	þyncan ‘to seem’	þuhte	geþuht
II	endian ‘to end’	endode	geendod
III	habban ‘to have’	hæfde	gehæfd
	secgan ‘to say’	sægde	gesægd
	hycgan ‘to think’	hogde	gehogod
	libban ‘to live’	lifde/leofde	gelifd/geleofd

Figure 3: Classification of Old English weak verbs.

The strong verbs in Old English are generally classified in seven classes, depending on the ablaut gradation they display in their citation forms, those being, the infinitive, the preterit singular, the preterit plural and the past participle. The traditional classification of these verbs is as follows in figure 4:

	Infinitive	**Preterit Singular**	**Preterit Plural**	**Past Participle**
I	drifan ‘to drive’	draf	drifon	gedrifen
II	cleofan ‘to cleave’	cleaf	clufon	geclofen
III	drincan ‘to drink	dranc	druncon	gedruncen
IV	beran ‘to bear’	bær	bæron	geboren
V	gifan ‘to give’	geaf	geafon	gegifen
VI	standan ‘to stand’	stod	stodon	gestanden
VII	slæpan ‘to sleep’	slep	slepon	geslæpen

Figure 4: Classification of Old English strong verbs (based on Pyles and Algeo 1982).

Figure 4 reflects the traditional basic classification of the strong verbs. However, further subdivisions have been proposed by several authors (Levin 1964; Lass 1994; Hogg 1992; Ringe 2006) on the basis of different vocalic developments explainable throughout historical or phonological factors. This and other questions regarding the different structures of the strong verb system will be reviewed in section 2.4.3.1 below.

Within the strong verbs, there is a particular group developed from the application of Verner’s Law. Old English does not permit /*h*/ to fall between voiced sounds; in this context, it is always dropped, and the preceding vowel is

lengthened. Thus, verbs like **seohan* turn out contracted forms like *sēon* ‘to see’. Contracted verbs are to be found only in classes I (*lēon* ‘to grant’), II (*wrēon* ‘to cover’), V (*sēon* ‘to see’), VI (*flēan* ‘flay’) and VII (*fōn* ‘to take’) of the strong verbs.

The conjugation of the present system of these verbs deserves some attention:

	Infinitive	
	sēon	
	Present	
1 sg.	sēo	
2 sg.	siehst	
3 sg.	siehþ	
Pl.	sēoþ	
	Subjunctive	
Sg.	sēo	
Pl.	sēon	
	Imperative	
Sg.	seoh	
Pl.	sēoþ	
	Inflected Infinitive	**Present Participle**
	sēon sēonne	seonde

Figure 5: The present system of the contracted verb *sēon* ‘to see’.

Apart from weak and strong verbs, in the Old English verbal system there were also verbs which did not accommodate in the previous groups. These irregular verbs included preterit present verbs and anomalous verbs.

Old English had a few verbs that were originally strong but whose strong preterit came to be used in a present-time sense; consequently, they had to form new weak preterits and to do so, they made use of the productive resource at that time, that is, the addition of a dental suffix, as in *sceal* ‘shall’, *sceolde* ‘should’. These verbs are called *preterit present verbs* and can be classified following the pattern of the strong verbs by considering their present form (originally a strong past) as seen in figure 6:

	Infinitive	**Present**	**Preterit**
I	agan ‘to owe’	āh	āhte
II	dēagan ‘to avail’	dēag	dohte
III	cunnan ‘to know’	cann	cūþe
IV	sculan ‘to be obliged’	sceal	sceolde
V			
VI	*motan	mōt	mōste
VII			
Uncertain	magan ‘to be able’	mæg (‘may’)	meahte (‘might’)

Figure 6: Preterit present verbs (based on Campbell 1987).

As can be seen in Figure 6, the preterit present verbs are the main source for the important group of modal verbs in Modern English, although these were not the only verbs included in this group. In fact, although originally not a part of this group in Old English, the verb *willan*, preterit *wolde* ‘to wish, want’ (the origin of Modern English *will, would)*, also became a part of the present-day modal system.

Finally, some of the most common verbs in Old English developed irregularities of different kinds. *Beon* ‘to be’ was in Old English, as its modern descendant combines alternative present indicative forms from several different roots, as follows:

(2)
(ic) *eom* or *beo* ‘I am’
(þu) *eart* or *bist* ‘you (sg.) are’
(he, heo, hit) *is* or *bið* ‘he, she, it is’
(we, ge, hi) *sindon, sind, sint,* or *beoð* ‘we, you, they are’

The forms *eom, is,* and *sind(on)* or *sint* were from an Indo-European root **es-*, with the forms **esmi, *esti,* and **senti,* seen in Sanskrit *asmi, asti,* and *santi* and in Latin *sum, est, sunt.* The second person *eart* was from a different Indo-European root, **er-* with the original meaning ‘arise’. The Modern English

plural *are* is from an Anglian form derived from that same root. The forms beginning with *b* were from a third root **bheu-*, from which came also Sanskrit *bhavati* 'becomes' and Latin *fui* 'have been'. The preterit forms were from yet another verb, whose infinitive in Old English was *wesan* (a Class V strong verb):

(3)
(ic) *wæs*
(þu) *wære*
(he, heo, hit) *wæs*
(we, ge, hi) *wæron*

The alternation of *s* and *r* in the preterit was the result of Verner's Law. The Old English verb for 'be', like its Modern English counterpart, combined forms of what were originally four different verbs (seen in the present-day forms *be, am, are, was).* Paradigms displaying historically unrelated forms are called *suppletive.*

Another suppletive verb is *gan* 'go', whose preterit *eode* was doubtless from the same Indo-European root as the Latin verb *eo* 'go'. Modern English has lost the *eode* preterit but has found a new suppletive form for *go* in *went,* the irregular preterit of *wend.*

Also irregular, and so anomalous, although not suppletive, is *don* 'do' with the weak preterit *dyde.*

Once the different classes of strong verbs have been introduced, section 2.3.3.1. raises the questions of why and how strong verbs are assigned to the different classes.

2.4.3.1. The classification of strong verbs

As has been pointed out in the description of the historical evolution of the strong verb system from Proto-Indo-European to Old English, the

classification of the strong verbs in the latter period very much depends on the perspective adopted. If we remain faithful to the Germanic distinction between ablauting and reduplicating verbs, only six classes of strong verbs are displayed. Thus, Hogg (1992):

	Present	**Pret. sing**	**Pret. Plural**	**Past. Part.**
I	ī	ā	i	i
II	ēo	ēa	u	o
III	e	æ	u	o
IV	e	æ	ǣ	e
V	e	æ	ǣ	o
VI	æ	ō	ō	æ

Figure 7: The ablauting verb classes in Old English (based on Hogg 1992).

Although of a different formative origin, the Indo-European reduplicating verbs have traditionally been included as the seventh class of strong verbs in Old English. These verbs display one vowel in the infinitive and the past participle, although it is not predictable, and another one in the preterit singular and plural, *e* or *eo*. Figure 8 represents this pattern:

	Present	**Pret. sing**	**Pret. Plural**	**Past. Part.**
VII	X	e/eo	e/eo	X

Figure 8: Old English Strong Class VII.

The vocalic contrasts of the different classes have been largely discussed, and some divergence has arisen as regards the (sub)classification of the non-standard verbs.

Thus, Class I remains homogeneous. Only Verner's Law causes the voicing of voiceless fricatives in the preterit plural and past participle forms. On the other hand we find preservation of Verner's Law, due to the tendency of Old English to the voicing of medial fricatives, "this is to be expected, for levelling of Verner's Law reduced allomorphic variation and simplified the system" (Hogg 1992: 152). From a phonological perspective (Hogg and Faulk 2011: 234), the ablaut grade of this class presents an ablaut pattern reflecting a

Proto-Indo-European ablauting vowel plus */j/. In West Saxon, back umlaut can be expected in the third principal part (preterit plural) of strong verbs of this class (e.g.: *belīfan* 'to remain', drīfan 'drive'), but it has been levelled, while back umlaut is common before nearly all consonants.

As regards, Class II Hogg (1992: 153) and Krygier (1994: 39) acknowledge the existence of three sub-types. Some class II verbs present *u* instead of *eo* in the first principal part like *scufan* 'to shove'. The other two subtypes present changes in the consonants and not in the root vowel. Some verbs present a change from *s* to *r* in the last two principal parts as *leosan* 'lose'- and *freosan* 'freeze' (*luron-geloren* and *fruron-gefrorern* respectively) while others show variation from *ð* to *d* in those *of seoðan* 'seethe,' which makes *sudon-gesoden*. On top of these changes, class II shows an ablaut pattern reflecting a Proto-Indo-European ablauting vowel plus */w/. In Anglian there is smoothing in the two principal parts when the stem ends in *c*, *g*, or *h*. This smoothed vowel can be present as well in the third principal part, so the result is that the plural preterit forms resemble those of Class V.

Class III verbs present several context dependent vocalic patterns. The vocalic alternations very much depended on the first element of the consonantal cluster that followed the root vowel. The standard pattern displayed in Figure 6 is used when the root vowel is followed by a nasal. If the consonant following the root vowel is an *l*, the gradation became *e, ea, u, o*, as is the case with *helpan-healp-hulpon-geholpen* 'to help'. If the consonant is *r* or *h*, the gradation was *eo, ea, u, o*: thus, *ceorfan-cearf-curfon-gecorfen* 'to carve' or *feohtan-feaht-fuhton-gefohten* 'to fight'.

Class IV has a short list of verbs, yet it displays a subdivision between those with a following liquid and those with a following nasal, namely *cuman* 'to come' and *niman* 'to take'. Those with the liquid were straightforward and followed the standard pattern. The other verbs behaved differently. In *cuman* only the preterit plural and the past participle are phonologically predictable (Hogg 1992: 154) and showed the pattern *u-ō-ō-u*. *Niman* displayed the series *i-a-ā-u*. These irregular series were the result of analogical formations and levelling with other verbal classes, namely Class III and VI.

Class V is very similar to class IV and its verbs are only distinguishable in the past participle. The difference is such because "the post-nuclear consonant is not a sonorant, and Proto-Indo-European reduced grade in the past participle is thus realised as /e/ rather than /o/" (Hogg and Faulk 2011: 246). Verner's Law could apply, as the consonant following the ablaut vowel could be any but a liquid consonant. However, this law only applied originally in the past participle, so the past plurals showing a modified consonant demonstrate the existence analogical extensions of this Law, as proved by the form *cwǣdon* 'they said' from the verb *cweðan*. This analogical extension indicates how closely related these two forms, and so they were treated similarly to their counterparts forms in classes I and II were Verner's Law fully applies.

Class VI should present a variation between /a/ and /æ/ in the present and past participle, but the standardisation of /a/ in the present and also in the past participle gave away with this dichotomy in some dialects, especially in West Saxon. Thus we get series like *faran-fōr-fōron-faren.*

Class VII, on its part, is the one that presents a greater number of sub-types. Krygier (1994) proposes seven different subtypes according to the following patterns:

(4)

a. ā-ē-ē-ā

b. ēa-ēo-ēo-ēa

c. ea-ēo-ēo-ea

d. aN-ē(o)N- ē(o)N-aN

e. ā/æ-ē(o)-ē(o)-ā/æ

f. ō-ēo-ēo-ō

g. ū-ø-ø-ū

Although this classification of the Old English strong verb system is generally accepted by most scholars, it is far from being undisputed. Levin (1964) argues that a purely synchronic approach, free from historical considerations would render a considerably different organisation. He pays attention not only to the

citation forms but also to the “differences of vowel grade in the remainder of their paradigms” (Levin 1964: 251). Furthermore, the traditional classification is organised following a reflection of Indo-European and Germanic root formations of different types. The criteria that are taken into account in the traditional analysis are “the particular distribution of resonants and consonants in the original root structure, the original ablaut alternation, and reduplication” (Levin 1964: 251). Thus, Class I reconstructs to a root with the resonant *i* followed by a consonant; class II to a root with *u* followed by a consonant; Class III to a root with a nasal or a liquid followed by a consonant. Class IV, on the other hand, reconstructs to a root ending in a nasal or a liquid resonant, and class V to a root ending in a consonant (stop or spirant); class VI is less determinate: the phoneme with which it ends is a stop, a spirant, a nasal, or a liquid. However, Levin states that between the reconstructed root structures of the Germanic Strong verbs and their correspondent forms in Old English there have been great changes. In classes I and II, the articulation of the resonants *i* and *u* with other elements of the root has been shaded by other sound-changes, among which Levin (1964: 253) includes the fact that “the preterit plural and the past participle continue to show the reconstructible form, but the present and the preterit singular do not”. In class III, there is little retention of the reconstructible forms; vowels which have suffered epenthesis have developed in the preterit plural and past participle to support the zero-grade nasals and liquids, and the original *e* and *o* of the present and the preterit singular have also undergone sound-change. Old English classes IV-VI keep the original consonant structure, but there are irregularities in the ablaut alternation.

Finally, Levin claims against the segregation of class VII and points out that “the historical basis for segregating the seventh class of Germanic strong verbs is that their preterits were originally formed by reduplication […] but only Gothic seems to have maintained this formation. In Old English, apart from some remainders, the seventh class does not present such formation.” (Levin 1964: 252). In fact, Class VII shows a big modification of the structure of the preterit because “the original reduplication has been, for all practical (i.e., synchronic) purposes, lost” (Levin 1964: 253).

From Levin's arguments against the traditional system, a new classification of the Old English strong verbs emerges. It still displays seven classes, but with a high degree of subcategorisation. Levin's (1964) proposal is summarised in Figure 9 below:

Class	**Subclass**	**Infinitive**	**Pret. Sg.**	**Pret. Pl**	**Past Part.**	**Sample Verb**
1	a	ī	ā	i	i	bīdan
	b	ēo	ā	i	i	wrēon
2	a	ēo	ēa	u	o	bēodan
	b	ū	ēa	u	o	brūcan
3	a	i	a	u	u	bindan
	b	e	ea	u	o	helpan
	c	eo	ea	u	o	weorpan
	d	u	ea	u	o	spurnan
	e	e	æ	u	o	stregdan
4	a	a	ō	ō	a	beran
	b	e	æ	ǣ	e	metan
	c	ēo	ēa	ǣ	e	sēon
	d	i	æ	ǣ	e	biddan
5	a	a	ō	ō	a	faran
	b	ēa	ō	ō	a	slēan
	c	e	ō	ō	a	hebban
	d	ie	ō	ō	ea	scieppan
	e	æ	ō	ō	æ	stæppan
	f	u	ō	ō	u	cuman
	g	i	ō	ō	u	niman
6	a	ā	ē	ē	ā	hātan
	b	ǣ	ē	ē	ǣ	lǣtan
	c	ō	ē	ē	a	fōn
7	a	a	ēo	ēo	a	bannan
	b	ea	ēo	ēo	ea	fealdan
	c	ā	ēo	ēo	ā	blāwan
	d	ēa	ēo	ēo	ēa	bēoton
	e	ō	ēo	ēo	ō	blōtan

Figure 9: Levin's (1964) classification of the Old English strong verbs.

Levin (1964) develops this classification by taking the vocalism of the preterit as the starting point. Thus, given the data in figure 9 an initial distinction between two groups of classes can be made. On the one hand, classes 1-3, where the vocalisms of the two preterit forms are different. On the other, classes 4-7, where the two preterit forms display one and the same vowel.

With Levin's (1964) classification, the preterit forms become distinctive. That holds true for classes 5 through 7 and also for the preterit singular of 3a and the preterit plural of 4b. Levin (1964) acknowledges that the vocalisms of the preterit are not sufficient for distinguishing all the verbs. When this is the case, Levin (1964) resorts to the root structure to explain the class distinctions. He states that "the historical changes affecting the root structure of Old English verbs have, as it happens, preserved a difference here. Verbs of class 3 have roots ending in a nasal or liquid followed by a consonant; verbs of classes 2 and 4 do not have this structure" (Levin 1964: 255). This makes the preterit singulars in class 3b-d distinguishable from those of class 4c, 3 and 2.

The root structure will help to distinguish class 3 forms from those of classes 2 and 4. Besides, 4a contains a liquid in the root, although class 2 does not.

Vowel *u* appears in the past participles of 3a, 5f, and 5g, and root structure is again decisive. When the vowel is *a* in classes 5a, 5b, 5c, 6c, and 7a, root structure is normally determinative. In the last two classes, *a* is followed by a geminate nasal (7a) or nasal plus consonant (6c).

Present stems are also distinctive, at least in class 1a. On the other hand, the *ū* of 2b is also distinctive, the *ēo* of 2a can be also found in 1b and 4c, but verbs in those classes are contracted verbs, whereas verbs in 2a are not.

In class 3, the vowel *eo* of 3c is distinctive; the *i* and *u* of 3a and 3d can be differentiated from their counterparts in 5g and 5f by root structure. The *e* of 3b and 3d is distinguished from the *e* of 4a and 4b in the same way, and from the *e* of 5c by the lack of root-geminate or -*ri*-.

Opposite to this is the case with classes 5d, 5e and 7b. Class 5a is similar to 7a in the sense that both have *a*. In 7a, *a* is always followed by a

geminate nasal or nasal plus consonant, while in 5a it only arises in *standan* ‘to stand’. *ēa* in 5b occurs in a contract form, and distinguished from the *ēa* of 7d.

The *ǣ* of 6b is distinctive, the *ā* of 7c is followed by *w*, but in 6a this does never happen. The *ō* in 6c occurs in contraction, but in 7e it does not happen.

By way of conclusion, Levin (1964) compares both systems, the traditional one and his own and remarks that his classification is more distinctive. In his own words, “the use of vowel gradation as the primary criterion, with root structure a subsidiary marker, greatly simplifies the system. By modifying the role of root structure as a class index, and by abandoning reduplication altogether, we achieve classes much more uniformly differentiated” (Levin 1964: 256). Finally his stronger criticism of the traditional system is the lack of motivation. Levin states that “the reclassification, by disregarding historical differences and focusing on synchronic evidence, presents a system of Old English strong verbs which is properly motivated; the result is neater and more adequate than the traditional scheme” (Levin 1964: 256).

Other authors present also divergent systems, although without any explanation of the criteria that underlie their proposal. Thus, Sweet (1896) present the seven-class structure shown in figure 10:

Class	**Infinitive**	**Pret. Sg.**	**Pret. Pl.**	**Past Part.**
Ia. Fall-group (eo-preterits)	ea	ēo	ēo	ea
Ib. Fall-group (e-preterits)	ā	ē	ē	ā
II. Shake-group	a	ō	ō	a
III. Bind-group	i	a	u	u
IV. Bear-group	e	æ	ǣ	o
V. Give-group	e ie	æ ea	ǣ ēa	e ie
VI. Shine-group	ī	ā	i	i

VII. Choose-group	ēo ū	ēa ēa	u u	o o

Figure 10: Sweet's classification of strong verbs.

To round off this section, Figure 11 shows a comparison between the classifications that Clark Hall (1896), Krygier (1994), Levin (1964) and Sweet (1896) have proposed. The starting point for this comparison is the set of verbs Levin (1964) proposes as examples for his classification. I have arranged the four proposals from the most standard (Krygier's) to the most divergent (Sweet's):

Sample Verb	**Krygier**	**Clark Hall**	**Levin**	**Sweet**
bidan	I	1	1a	6
wreon	I contract	1_2	1b	6,7
seon	I contract	5	4c	7
bēodan	II	2	2a	7
brucan	II	2	2b	7
bindan	IIIa	3	3a	3
helpan	IIIb	3	3b	3
weorpan	IIIb	3	3c	3
spurnan	IIIb	3	3d	ø
stregdan	IIIc	3	3e	3
beran	IV	4	4a	4
cuman	IV	4	5f	4
niman	IV	4	5g	4
metan	V	5	4b	5
biddan	V	5	4d	5
faran	VI	6	5a	2
hebban	VI	6	5c	2
scieppan	VI	6	5d	2
stæppan	VI	6	5e	2
slean	VI contract	6	5b	2

hatan	VIIa	7	6a	1b
beatan	VIIb	7	7d	1
fealdan	VIIc	7	7b	1
bannan	VIId	7	7a	1a, b
fon	VIId contract	7	6c	ø
lætan	VIIe	7	6b	1b
blawan	VIIe	7	7c	1
blotan	VIIf	7	7e	1

Figure 11: A comparison of different strong verb classifications.

In figure 11, I have introduced Clark Hall's (1896) dictionary, for it is the dictionary upon which the database *Nerthus* was originally based. As can be seen, his classification is quite standard, but does not include any degree of subdivision. Levin's (1964) recategorisation affects basically to the traditional classes IV-VII, which are thoroughly subsumed and redistributed. Finally, Sweet's (1966) proposal is presented in a way that is practically a reversal of the traditional classification.

For the purposes of this work, I will draw on Krygier's (1994) proposal as it is usually accepted as the standard classification. It gives way to variation and distinctions class-internally and also accounts for the distinctive contracted verbs. Once the different classifications of strong verbs have been discussed, the following section will deal with the inflectional features of the strong verbs, as well as with their spelling variations resulting from morphophonological or historical factors.

2.4.3.2. The inflectional paradigm of strong verbs and spelling variations

This section focuses on the analysis of the strong verb inflectional paradigm, its conjugation and main characteristics by considering dialectal variants and phonological alternations which are crucial for a study of verb forms and alternant forms and spellings, like the one undertaken in this research.

As remarked above, Old English distinguishes two voices, although the passive is basically periphrastic, so most verbal forms belong to the active voice. As regards moods, there is an indicative and a subjunctive mood, with two tenses each, present and past, along with an imperative mood, with just a present tense, and an infinitive mood. The indicative and the subjunctive moods distinguish three persons in the singular, and only one in the plural. It must be noted though, that the singular forms of the subjunctive are identical, and so usually subsumed under the label *singular* without further person distinction.

The different verb forms were characterised by the inflectional endings. The infinitive is signalled with an *-an* ending (*cēowan* 'to chew'), with the exception of the contracted verbs where only the *-n* is preserved (*sēon* 'see', *fōn* 'to take'). In the present, the first person singular is marked by an *-e* ending (*cēowe*), whereas the second person is signalled by the ending *-st* (*cīewst*) and the third is marked by *-þ* (*cīewþ*). The plural form was indicated by means of the ending *-aþ* (*cēowaþ*). In the past tense, along with the ablaut grade, there was a zero ending in the first and third persons of the singular (*cēaw*), while the second one was marked by the ending *-e* (*cuwe*). The plural form was distinguished by the traditional *-on* (*cuwon*) ending. The subjunctive mood displays an *-e* ending for the singular and an *-en* ending for the plural both in the present and past tenses (cēowe, cēowen, cuwe, cuwen). Apart from the infinitive, Old English made use of an inflected infinitive, a reminiscence of earlier times. It was used especially when the infinitive had a nominal function, and shows an *-enne* ending (*to cēowenne*). Finally, there were two participles, present and past, whose endings were *-ende* for the present participle (*cēowende*) and *-en* for the past one. Besides its ending, the past participle was also indicated by the addition of the prefix *ge-* (*gecowen*). The imperative mood distinguishes a singular form with no ending (*cēow*), and a plural form with the ending *-aþ* (*cēowaþ*), which coincides with the indicative present plural form.

These endings did not only account for the distinction of the different verb forms. They may also have consequences on the phonological and

spelling levels. That is the case with the endings of the second and third person singular indicative derived from the Germanic *-isi, -iþi* forms. They produced a standardised development in Old English through *i*-umlaut by means of which the root vowel *e* changes into *i* and the diphthong *ēo* into *īe*. This was so even if the initial *-i-* of the inflectional ending is syncopated. Campbell (1987: 299) acknowledges that in West Saxon "there is generally syncope of *-i-* and consequent assimilation of consonants and simplification of double consonants within groups".

In this vein, Wright (1984) remarks that in the second and third person singular the *-i-* (*-e-*) was syncopated after long stems (*hilpst, tīehst*), and remained after short stems (*birest, færest*). There are exceptions to this rule, especially in West Saxon and Kentish, and so, new formations were created in both directions (*bindest, bindeþ*; and *birst, birþ*). In Anglian, forms without syncope were generalised, but in West Saxon and Kentish syncope was general, especially after voiceless consonants and after d, f and g; but not after a single liquid or a nasal. Campbell (1987) adds that in the Northumbrian dialect syncope hardly ever takes place, and the mutation of the root vowel is levelled away. The third singular forms in early Kentish are not contracted but there is prevailing syncope, umlaut and consonant assimilation, while the root vowel *e* is restored. Finally, in West Saxon, some Anglian-type forms occur. Hogg and Faulk (2011) also insist on this feature and state that, as a rule, syncope of the high vowel in the inflexion occurs after a heavy syllable, and not so often after a light one. "Amongst the heavy stems, the chief exceptions are those verbs (all weak) that have a stem ending in a post-consonantal sonorant […] otherwise, when syncope does not occur after heavy stems in West Saxon, the cause may be stylistic, or the text containing such forms may be of Anglian origin" (Hogg and Faulk 2011: 217).

However, these two endings were also significant, not only because there was influence upon the root vowel of the verb, but because of the consonantal developments that originated when the initial *-i-* of the ending was lost. Under those circumstances, groups of consonants were created which were subject to further modifications. These modifications (both partial and

total assimilations) are exemplified and explained by Campbell (1987: 299-300), who puts forward the following examples for the 2nd person singular of the present indicative.

> -When a voiced dental plosive gets in contact with the inflectional ending, the plosive is devoiced, *dst>tst* as in *rǣtst* (from *rǣdan* 'to decide'); under the new context *-t* can be dropped as in *finst* (*findan* 'to find'). The *-t* is also syncopated even if it is there originally and not as a result of assimilation as in *hǣst* (*hātan* 'to call').
>
> -Total assimilation takes place when the elision of the vowel brings a voiceless dental fricative in contact with the inflectional ending (*þst>sst*). In this case, the double consonant is simplified (*sst>st*) as in *cwist* (*cweþan* 'to say'). This process is later on reversed, and the original *þ* is restored by analogy as happens in *snīþst, snītst* (*snīþan* 'to cut').
>
> -In the cluster *-ngst* two phenomena do occur. In the first place, there is devoicing of the plosive by means of partial regressive assimilation, as in *brincst* (*bringan* 'to bring'). It may also be the case that the plosive is completely lost, giving way to forms like *brinst.*
>
> -In the group *-gst*, when not preceded by *n*, *g* is devoiced and becomes *h* thus creating the group *-hst* and consequently forms like *stīhst* (*stīgan* 'to move')
>
> -Finally, the group *-hst* may become *-xt* though this is rare, and happens only in the West Saxon dialect. Attested forms include *syxt* instead of *sīehst* (*sēon* 'to see').

In the 3rd person singular indicative, there are four consonantal changes due to the influence of the inflectional ending:

> - When the inflectional *-þ* comes into contact with a dental plosive the resulting form is *-tt*. So, if the plosive is voiceless /t/, it produces total progressive assimilation of the inflectional ending, as in *lǣtþ>lǣtt*

(*lǣtan* ‘to allow to remain’). But when the plosive is voiced, then both consonants have influence upon each other. The inflectional *þ* causes the devoicing of the plosive *dþ>tþ* and in that context, there is assimilation as in the case of an original *t* as in *bīedþ>bīetþ>bīett* (*bēodan* ‘to command’).

-In contact with a voiceless alveolar fricative *s* the inflectional *þ* becomes a plosive. Thus, *sþ>st* as in *cīest* (*cēosan* ‘to choose’).

- The group *gþ* undergoes devoicing of its initial member, which moreover, becomes fricativised into *h* as in *līehþ, abelhð, byrhð* (*lēogan* ‘to lie’, *belgan* ‘to be or become angry’, *beorgan* ‘to save’)

-Finally, as happens with the second person, in the group *ngþ* there is devoicing of the medial plosive, thus turning out *ncþ* as in *brincþ* (*bringan* ‘to bring’).

These general tendencies may be overcome by dialectal variation. Usually, the second and third person singular of the present indicative are not syncopated in Anglian texts as shown by the forms *gehifð, gefiht* instead of *gefið*, or *fōeht* instead of *fōeð*. In Northern Mercia, syncope is not normally present, and umlaut does not take place. Consider the forms *falleð, ceorfeð, weorðeð, cēoseð, lūceð, scadeð, cnāweð, flōweð*.

In this region, *i* is retained in the root syllable as in *trides, ites, ætfileð, bireð, gildeð, swilteð*. (from *tredan* ‘to tread’ and *etan* ‘to eat’). Contrasted to this, in the rest of Mercia, *-e* is restored as *-a*.

The third person singular forms appearing in east Kentish are uncontracted, as is the case of *bibēadeð, forgifeð*; but there is syncope, umlaut and consonant assimilation. Examples of this are *gebēgð, toīot, helt*, inf. *būgan* ‘to bow’, *gēotan* ‘to pour’, *healdan* ‘to hold’. The root vowel *e* is often restored, as in *gelpð, tret* (*gelpan, tredan* ‘to tread’) and so, the second singular *byrst* (*beran* ‘to bear’) where *-y* stands for *-e* in Kentish or *ægelts* (*geldan*). The vowel of the third person singular in *-et* is not syncopated: *forberet, aweget*, inf. *beran* ‘to bear’, *wegan* ‘to carry’. rule, as in *eteþ* (*etan* ‘to eat’ or *swælteþ* (sweltan ‘to die’).

Other generalisations regarding verbal dialectal variation of the inflectional forms include:

- The first singular present indicative appears in two forms, *-e*/*-æ* and *-o*/*-u*. *-E* is practically universal in West Saxon, the only exception in early West Saxon being *cweðo*. The West Saxon present indicative ending *-e* in of disputed origin, "but it is usually regarded as borrowed from the subjunctive" (Hogg and Faulk 2011: 217). Early Kentish has usually *-e*, but there are instances of *-o*, as in *hāto* 'command', *biddo* 'bid', maybe under Mercian influence. Early Anglian texts have prevailingly *-o*/*-u*, with a few cases with *-e*. In late Northumbrian *-o* prevails, *-a* is frequent, *-u* is less frequent, and *-e* appears quite often. In Anglian, monosyllabic forms occasionally add *-m* on the analogy of *eam* ' I am'. The differences on the inflexion of the first person singular of the present indicative in Anglian, West Saxon and Kentish, keeps with developments in the second and third person singular of the same verbal tense.
- The second and third singular present indicative have *-i-* as the vowel of the endings in early texts. Forms with *-i-* will occur later sporadically in Northumbrian. In earlier times, the second singular ends in *-s*, but later, *-st* replaces this widely. Early West Saxon has always *-st*, or *-sð*; Northumbrian has *-st* rarely, normally only in syllabic forms. The third singular has *-it*, *-et* beside *-iþ*, *-eþ*. Some documents regard it as graphic, but the appearance of the form in later texts suggests a truly phonetic variant. In early West Saxon all the examples are from weak verbs *gescīnet, fallet*.
- The plural of the present indicative shows little variation outside Northumbrian. In Northumbrian, *-as* is more frequent than *-að*, and in both forms the reduction of the vowel is frequent, giving *-eð*, *-es*.
- In the preterit indicative, the second person singular ending *-e* reflects the Germanic ending *-i*, which may have been lost after heavy syllables. Probably it was restored by analogy to the ending retained after light syllables. The second singular of the past indicative sometimes adds -

s(t) in Northumbrian, owing to analogy of the weak verbs. Under those circumstances, Old English has regularised the ending of the second person singular into *-þ* although the examples are limited to verbs of Class VII with root syllables ending in *-t*.

- The plural of the preterit indicative has *-un* in early texts, beside *-on*; the former variant is common in early Mercian texts. The variant *-an* is common in early West Saxon, and *-en* is found occasionally in Anglian texts. In early Northumbrian, there are cases of preterit plurals in *-u*; *-n* is rarely lost in late Northumbrian, which has *-on* beside *-un*.

- In the subjunctive, the distinction between indicative and subjunctive forms is maintained in early West Saxon, but in late West Saxon the subjunctive plural form *-on* comes from the indicative mood, both in the present and the preterit, and *-on*, *-an* appear in early West Saxon in the preterit. In Northumbrian plural inflexions are not differentiated due to the loss of the final *-n*, so the plural inflexions are *-e*, *-a*, *-æ*, but *-en* is not very common. The preterit subjunctive in Old English made use of the old endings of the present subjunctive before the *i*-umlaut took place and consequently, there is absence of *i*-umlaut in these forms.

- The imperative forms were also subject to modification. The third person plural of the present indicative was used for the second person plural imperative, while a form of *-an* was used in the oldest period of the language for the first person plural. “This form was originally identical with the first person plural of the present indicative which disappeared in OE” (Wright 1984: 259).

- The infinitive underwent modifications in both its forms. The suffix *-an* derived from the primitive Germanic *-onom*, and it was syncopated in Northumbrian being the final *-n* lost. It appears in all early Mercian and southern texts in *-an*. *-on* appears later in late West Saxon. As stated, in early Northumbrian, the infinitive occurs with loss of *-n*. Late Northumbrian has always loss of *-n*, the frequent vowel is *-a*, but *-e* is not uncommon, and it is possible to find *-æ* and *-o*. On the other hand,

the inflected infinitive *-enne* evolved into *-anne* through the influence of the *-an* infinitive ending.

- The ending of the present participle is *-ende* in all texts, although in Northumbrian occasional *-ænde* is regarded as a spelling variant. This forms evolved from *-ændi*, to *-endi* and later to *-ende*.
- The past participle forms of primitive Germanic *-énaz*, *-íniz* were regularised in Old English into *-en*, but not in the oldest period of the language where *æn*, *-en* and later *-in* were preferred.

As has been shown, though much simpler than its cognate and mother tongues, Old English is still subject to some degree of inflectional variation partly attributable to regional differences.

Leaving aside inflection, Old English is characterised by generalised word-formation. Old English made use of the native words to develop new lexical items through affixation (both prefixation and suffixation), compounding, and zero-derivation. The following subsection will engage in the description of these phenomena, with the focus put on derivations based upon the strong verb.

2.4.3.3. The derivation based on the strong verb

Hinderling (1967), Bammesberger (1965), Seebold (1970), Kastovsky (1992, 2006) and Martín Arista (2010, 2011a, 2011b, 2012, 2013) among others, consider the strong verb the starting point of lexical derivation in Germanic in general and Old English in particular. Heidermanns (1993) as well as Wodtko, Irslinger and Scheneider (2008), while focusing, respectively, on adjectives and nouns, also acknowledge the central role played by the strong verb.

In Old English, lexical derivation from strong verbs can take place by zero derivation (without explicit derivational morphemes), affixation (prefixation or suffixation of free forms) or compounding (combination of free forms).

Previous studies in the zero derivation of Old English include Bammesberger (1965), Jensen (1913), Kastovsky (1968, 1992, 2006), Martín Arista (2011a), Palmgren (1904) and Schuldt (1905). Zero derivatives can belong to two types depending on the morphological contrast between base and derivative. Some zero derivatives display one of the root vowels of the ablaut pattern of the strong verb, that is, the infinitive, the preterit singular, the preterit plural and the past participle. Other zero derivatives hold a vocalic contrast with the strong verb that is recurrent and systematic and is called *alternation*.

Let us consider, in the first place, zero derivatives based on the ablaut pattern of the strong verb. Within the class of nouns, Palmgren (1904) relates some nominal derivatives to the infinitive of strong verbs, as is the case with feminines like *dræge* 'drag-net' ~ *dragan* 'to drag', masculines like *bēod* 'table, bowl' ~ *bēodan* 'to offer' and neuters like *geðeot* 'howling' ~ *ðeotan* 'to howl'. Nominal derivatives can also display the same vowel as the preterit singular, thus the feminine *tēag* 'chain' ~ *tēon* 'to draw', the masculine *swam* 'fungus' ~ *swimman* 'to swimm'. Nominal derivatives can also share the root vowel with the preterit plural, as in the feminine *wǣge* 'scales, balance' ~ *wegan* 'to weigh' and the neuter *gestæn* 'groaning' ~ *stenan* 'groan'. The past participle shares its vowel with feminines like *storfe* 'flesh of animals that have died by disease' ~ *steorfan* 'to die', masculines like *blice* 'laying bare (of bone through wound)' ~ *blīcan* 'shine, be laid bare (of bone)' and neuters like *geðwit* 'what is shaved off, chip' ~ *ðwītan* 'cut, shave off'. In the class of the weak verb, Schuldt (1905) relates the root vowel of some weak verbs of class 1 to the preterit singular of the strong verb (*bǣtan* 'to bridle' ~ *bītan* 'to bite'). Other weak verbs display the vowel of the infinitive but show an alternation with the vowel of the preterit, as in *missan* 'to miss, fail' ~ *mīðan* 'to conceal'. Some class 2 weak verbs share the vowel of the preterit singular of the strong verb, as is the case with *wracian* 'to be in exile' ~ *wrecan* 'to drive, press'. Other weak verbs from class 2 present a contrast of quantity with the past participle of the strong verb, such as *cunnian* 'to prove, try' ~ *cunnan* 'to be or become acquainted with'. Finally, weak verbs of class 2 can also be derived with the vowel of the present or the preterit plural of the strong verb, as is the

case with *treddian* ‘to tread, step, walk’ ~ *tredan* ‘to tread’. In the class of the adjective, Jensen (1913) identifies adjectives formed with the vowel of the present (*scīn* ‘extraordinary appearance’ ~ *scīnan* ‘to shine’ and *reōd* ‘red’ ~ *reōdan* ‘to redden’), with the vowel of the preterit singular (*hnāg* ‘bent’ ~ *hnīgan* ‘to bend down’ and *reāfol* ‘rapacious’ ~ *reāf* ‘to spoil’) and with the vowel of the past participle (*ǣswīc* ‘offensive’ ~ *swīcan* ‘to move about’ and *flugol* ‘apt to fly’~ *fleōgan* ‘to fly’).

Regarding zero derivatives with root vowels that hold an alternation with the vowel of the strong verb, Kastovsky (1968: 59) identifies the correspondences between nouns and verbs that are shown in figure 1 (see also Kastovsky 1986, 1989). Alternations can be broken down into vocalic (A) and consonantal (C). Vocalic alternations, in turn, fall into two types, direct and reverse (R). Direct alternations involve mainly strong verbs, while reverse alternations occur with respect to weak verbs in most cases.

A1	/a/ ~ /æ/	*faran ~ fær*
A1R	/æ/ ~ /a/	*stæl ~ stalu*
A2	/a/ ~ /e/	*acan ~ ece*
A2R	/e/ ~ /a/	*sendan ~ sand*
A3	/ea/ ~ /ie/	*feallan ~ fiell*
A3R	/ie/ ~ /ea/	*mierran ~ gemearr*
A4a	/e/ ~ /i/	*gecweden ~ cwide*
A4b	/eo/ ~ /ie/	*weorpan ~ wierp*
A4bR	/y/ ~ /eo/	*wyrcan ~ weorc*
A5	/o/ ~ /y/	*gebrocen ~ bryce*
A5R	/y/ ~ /o/	*spyrian ~ spor*
A6	/u/ ~ /y/	*burston ~ byrst*
A6R	/y/ ~ /u/	*hyscan ~ husc*
A7	/ā/ ~ /ǣ/	*drāf ~ drǣf*
A7R	/ǣ/ ~ /ā/	*lǣran-lār*
A8	/ō/ ~ /ē/	*lōcian ~ lēc*
A8R	/ē/ ~ /ō/	*fēdan ~ fōda*

A9	/ēa/ ~ /īe/	*hlēat ~ hlīet*
A9R	/īe/ ~ /ēa/	*īecan ~ ēaca*
A10	/ēo/ ~ /īe/	*flēotan ~ flīeta*
A10R	/īe/ ~ /ēo/	*stīeran ~ stēora*
A11	/ū/ ~ /ȳ/	*būan ~ bȳ*
C1	[non-palatal] ~ [palatal]	*acan ~ ece*
C1R	[palatal] ~ [non-palatal]	*þencan ~ þanc*
C2	/short consonant/ ~ /long consonant/	*tredan ~ tredde*
C2R	/long consonant/ ~ /short consonant/	*sellan ~ sala*
C3	/k/ ~ [x]	*tǣcan ~ getāh*
C4	[g] ~ /j/	*dragan ~ dræge*
C4R	/j/ ~ [g]	*wegan ~ wegu*
C5	[q] ~ /d/	*scrāð ~ scrād*
C6	/d/ ~ [q]	*scridon ~ scriðe*
C7	[voiced fricative] ~ [voiceless fricative]	*delfan ~ delf*

Figure 12: Verb-noun alternations (Kastovsky 1968).

Pilch (1970) identifies zero derivatives from strong verbs of the lexical classes of the noun, the verb and the adjective. Within the nominal class, we find in the first place strong masculine/neuter nouns without ending in the nominative singular from strong verbs and weak verbs of the class 1: *wīg* ‘strife, war’ n. < *wīgan* ‘to fight, make war’ *gielp* ‘boasting, pride, arrogance’ m. < *gielpan* ‘to boast, exult’. Some of them show gemination (*webb* ‘web’ n. < *wefan* ‘to weave’; *witt* m. ‘philosopher, wise man’ < *witan* ‘to know’) or prefixation with *ge-* (*gebind* ‘binding’ n. < *bindan* ‘to tie’; *gefeoht* ‘action of fighting’ n. < *feohtan* ‘to fight’. There is also a group of masculine and feminine nouns without ending in the nominative singular from strong verbs of classes IV, V and VI: *bǣr* f. ‘bier’ < *beran* ‘to bear’ str. IV; *æt* m. ‘food’ < *etan* ‘to eat’ str. V; *fōr* f. ‘going, course, journey’ < *faran* ‘to travel’ str. VI. Another group of nominal zero derivatives consists of strong masculine and feminine nouns with zero-stem in the nominative singular from strong verbs of classes I-VI. Instances based on strong classes include *wyrp* ‘a throw, cast’ < *weorpan* ‘to throw’ str. III, *rād* ‘ride, riding, expedition’ < *rīdan* ‘to ride’ str. I and *lād*

‘course, journey’ < *līðan* ‘to go, travel’ str. II. There is also a group of deverbal nouns from strong classes I-VI with zero in the nominative singular, including: *bite* ‘bite’ < *bītan* ‘to bite’ str. I, *lyre* ‘loss’ < *lēosan* ‘to lose’ str. II, *bryne* ‘corslet, burne’ < *biernan* ‘to burn’ str. III, *myne* ‘memory’ < *munan* ‘to think about, remember’ str. IV, *bryce* ‘breach’ < *brecan* ‘to break’ str. V and *hefe* ‘weight’ < *hebban* ‘to heave, raise’ str. VI. Also deverbal are strong feminine nouns with zero-stem in the nominative singular, from strong verbs of classes I-VI and weak verbs of classes 1 and 3 like *notu* ‘enjoyment, use’ < *nēotan* ‘to use’ str. II, *sagu* ‘saw, saying, report’ < *secgan* ‘to say, speak’ wk. III, *cwalu* ‘killing, murder’ < *cwelan* ‘to kill, murder’ str. IV, *wracu* ‘revenge’ < *wrecan* ‘to drive, impel, revenge’ str. V and *faru* ‘way, going, journey’ < *faran* ‘to set forth to, travel’ str. VI. Deverbal agentives of the weak declension constitute an outstanding group of zero derived nominals. They are based on verbs from the weak classes 1 and 2 like *dēma* ‘judge, ruler’ < *dēman* ‘to judge’ wk. 1 and *hunta* ‘huntsman’ < *huntian* ‘hunt’ wk. 2; and on strong verbs of classes I-VII, as in instances like *wita* ‘sage, philosopher’ < *witan* ‘to know, understand’ str. I, *boda* ‘messenger, herald, angel’ < *bēodan* ‘to command; inform’ str. II, *fricca* ‘herald, crier’ < *frignan* ‘to ask, inquire’ str. III, *bora* ‘ruler’ < *beran* ‘to bear’ str. IV, *lida* ‘sailor’ < *līðan* ‘to go to, travel’ str. V, *stapa* ‘grasshopper’ < *steppan* ‘to step’ str. VI and *wealda* ‘ruler’ < *wealdan* ‘to rule’ str. VII.

In the class of the adjective, Pilch (1970) finds deverbal adjectives without derivational suffix ending in strong masculine singular *-e* and formed from strong verbs like *swice* ‘fallacious, deceitful’ < *swīcan* ‘to wander’ str. I, *lyge* ‘lying, false’ < *lēogan* ‘to lie’ str. II, *gefrǣge* ‘well-known’ < *gefrignan* ‘to ask, inquiry’ str. III, *bryce* ‘fragile’ < *brecan* ‘to break’ str. IV, *ungemete* ‘huge’ < *metan* ‘to measure’ str. V, *gefēre* ‘accesible’ < *faran* ‘to set forth’ str. VI and *oncnǣwe* ‘known, recognised’ < *cnawan* ‘to know’ str. VII. Schön (1905) provides a list of adjectives zero derived from strong verbs that includes *reod*: ‘colour red’ < *rēodan*: ‘to redden, stain with blood’ str. II, *seoc*: ‘sick, ill’ < *sēocan*: ‘to be ill, fall ill’ str. II, *(ge)sceot*: ‘ready, quick’ < *scēotan*: ‘to shoot’ str. II, *ǣswind*: ‘idle, slothful’ < *swindan*: ‘to vanish,

consume’ str. III, *grimm*: ‘fierce, savage’ < *grimman*: ‘to rage, hasten on’ str. III, *melc*: ‘giving milk, milch’ < *melcan*: ‘to milk’ str. III, *sciell*: ‘sonorous, shrill’ < *sciellan*: ‘to sound’ str. III, *unmurn*: ‘untroubled’ < *murnan*: ‘to care, be anxious or fearful about’ str. with weak forms, *sac*: ‘accused, charged, guilty’ < *sīcan*: ‘to struggle, dispute, disagree, wrangle, fight’ str. I, *gescād*: ‘resonable, prudent’ < *sceadan*: ‘to divide, separate, part’ str. VII, *anweald*: ‘powerful’ < *wealdan*: ‘to rule, control, determine, direct, command, to have power over’ str. VII, *blāc*: ‘bright, shining, glittering, flashing’ < *blīcan*: ‘to glitter, shine, gleam, sparkle, dazzle’ str. I, *ðān*: ‘moist, irrigated’ < *þinan*: ‘to grow moist’ str. with weak forms, *wāc*: ‘yielding, not rigid, pliant, fluid’ < *wīcan*: ‘to yield, give away, fall down’ str. I, *hnāg*: ‘bent down, abject, poor, humble, lowly’ < *hnīgan*: ‘to bow oneself, bend, bow down’ str. I, *read*: ‘red (of gold)’ < *rēodan*: ‘to redden, stain with blood’ str. II, *geap*: ‘open, wide, extensive, broad, spacious, lofty, steep’ < *gēopan*: ‘to take in’ str. II, *unbeceas*: ‘incontestable’ < *cēosan*: ‘to choose, seek out, select’ str. II, *gram*: ‘angry, cruel, fierce’ < *grimman*: ‘to rage’ str. III, *scearp*: ‘sharp, pointed, prickly’ < *sceorpan*: ‘to scrape, to irritate’ str. III, *cōl*: ‘cool, cold’ < *calan*: ‘to grow cool or cold’ str. VI, *smolt*: ‘mild, peaceful, still, gentle < *gesmyltan*: ‘to appease, quiet’ str. with weak forms, *gedwol*: ‘heretical’ < *dwelan*: ‘to be led into error, err’ str. IV, *gemun*: ‘mindful, remembering’ < *munan*: ‘to think about, be mindful of, remember, mention’ str. with weak forms, *genōg*: ‘enough, sufficient, abundant’ < *genugan*: ‘to suffice, not to lack’ str. with weak forms, *lætt*: ‘slow’ < *latian*: ‘to be slow, indolent’ str. with weak forms, *brȳce*: ‘fragile, brittle, worthless, fleeting’ < *brecan*: ‘to break, shatter, burst, tear’ str. IV, *trede*: ‘fit to tread on, firm’ < *tredan*: ‘to tread, step on, trample’ str. V, *andfenge*: ‘acceptable, agreeable, approved, fit, suitable’ < *onfōn*: ‘to take, receive, accept’ str. VII, *ǣlǣte*: ‘desert, empty’ < *lǣtan*: ‘to allow to remain, leave behind, depart from, let’ str. VII, *gecnǣwe*: ‘conscious of, acknowledging’ < *cnāwan*: ‘to know, perceive’ str. VII, *genge*: ‘prevailing, effectual, appropriate’ < *gangan*: ‘to go, walk, turn out’ str. VII, *(ge)rǣde*: ‘prepared, ready, ready for riding (horse)’ < *rǣdan*: ‘to advise, counsel, persuade’ str. with weak forms, *gestence*: ‘odoriferous’ < *stincan*: ‘to emit a

smell, stink, exhale' str. III, *gescrence*: 'withered, dry' < *scrincan*: 'to shrink, contract, shrivel up, wither, pine away' str. III, *fēre*: 'able to go, fit for (military) service' < *faran*: 'to set forth, go, travel, wander, proceed' str. VI, *swice*: 'fallacious, deceitful' < *swīcan*: 'to wander' str. I, *hryre*: 'fall, descent, ruin, destruction, decay' < *hrēosan*: 'to fall, sink, fall down, go to ruin' str. II, *lyge*: 'lying, false' < *lēogan*: 'to lie' str. II, *nytt*: 'useful, beneficial, helpful, profitable' < *nēotan*: 'to use, have the use of, enjoy, employ' str. II, *twimylt(e)*: 'twice-melted' < *meltan*: 'to consume by fire, melt, burn up' str. III, *scrynce*: 'withered' < *scrincan*: 'to shrink, contract, shrivel up, wither, pine away' str. III, *earfoðfynde*: 'hard to find' < *findan*: 'to find, meet with' str. III, *gemyne*: 'mindful' < *munan*: 'to think about, be mindful of, remember' str. with weak forms, *andgiete*: 'understanding, intellect' < *gietan*: 'to get' str. V, *gesprǣce*: 'eloquent, affable' < *sprecan*: 'to speak, say, utter, make a speech' str. V, *gefrǣge*: 'well-known, celebrated, reputable' < *frignan*: 'to ask, inquire' str. III, *unbrǣce*: 'unbreakable, indestructible' < *brecan*: 'to break, shatter, brust, tear' str. IV, *micelǣte*: 'greedy' < *etan*: 'to eat' str. V, *mǣte*: 'mean, moderate, poor, inferior, small, bad' < *metan*: 'to meet, find, find out, fall in with, encounter' str. V, *bryce*: 'breach, fracture, breaking, infringement' < *brecan*: 'to break, shatter, burst, tear' str. IV.

Considering the verbal class, strong verbs do not produce new strong verbs. There is a group of weak verbs derived from strong verbs with o-stem and i-umlaut in the root vowel, including *rǣran* (<*raizjan)* 'to rear, raise' < *rīsani* 'to rise, stand up', *āflīegan* (<*flaugjan*) 'to put to flight' < *flēon (</fliuhan/)* 'to fly from, flee', *drencan* (<*drankjan*) 'to give to drink' < *drincan* 'to drink', *sengan* 'to singe, burn slightly' (*sangjan*) < *singan* 'to sing, sound', *swebban* (**swavjan*) 'to put to sleep' < *swefan* 'to sleep, slumber', *settan* (*satjan*) 'to make to sit' < *sittan* 'to sit, sit down'. Schuldt (1907) classifies the weak verbs derived from strong verbs without alternation on the grounds of morphological class. Weak verbs (class 1) derived from strong verbs without alternation with respect to the vowel of the third person singular present form of the strong verb include *missan* 'to miss, fail' < *mīðan* 'to conceal' str. I, *gebryttan* 'to break to pieces' < *brēotan* 'to bruise, break,

demolish' str. II, *brygdan* 'to turn' < *bregdan* 'to move to and fro' str. III, *geberian* 'to happen' < *beran* 'to bear, carry' str. IV and *ferian* 'to carry, convey' < *faran* 'to go' str. V. Among the weak verbs (class 2) derived from strong verbs with the vowel of the present or the preterit plural of the strong verb we find *spiwian* 'to spit up' < *spīwan* 'to spew, vomit, spit up' str. I, *ācreōpian* 'to creep' ~ *creōpan* 'to creep, crawl' str. II, *treddian* 'to tread, step, walk' ~ *tredan* 'to tread' str. V, *sacian* 'to strive' ~ *sacan* 'to fight, strive' str. VI and *hangian* 'to hang, be suspended' ~ *hōn* 'to hang, be suspended' str. VII.

As has been said above, the derivation from strong verbs takes place by means of zero derivation, affixation and compounding. Affixation, as is well known, comprises prefixation and suffixation. There is a long tradition of studies in the word-formation of Old English and, especially, in affixation. Previous work in prefixation includes, among others, Brinton (1986), Brinton and Traugott (2005), de la Cruz (1973), Hendrickson (1948), Hiltunen (1983), Hohenstein (1912), Horgan (1980), Lenze (1909), Lindemann (1953, 1970), Lüngen (1911), Martín Arista (2006, 2010a, 2010b, 2012), Meroney (1943), Mitchell (1978), Niwa (1966), Pilch (1953), Roberts (1980), Röhling (1914), Samuels (1949), Schrader (1913), Siemerling (1909), Trips (2009) and Weick (1911). Previous research in suffixation includes, among others, Bauer (2007), García García (2012a, 2012b), González Torres (2009, 2010a, 2010b, 2011), Hallander (1966), Haselow (2010), Kastovsky (1971), Lacalle Palacios (2011), Maíz Villalta (2011, 2012), Marckwardt (1942), Mateo Mendaza (2012, 2014), Nicolai (1907), Schabram (1970), Stark (1982), Torre Alonso (2010, 2011a, 2011b, 2012), von Lindheim (1958) and Weyhe (1910).

Beginning with the prefixation with strong verb bases, Kastovsky (1992), Lass (1994) and Quirk and Wrenn (1994) provide an inventory of prefixes that attach to strong verbs to form other strong verbs. Among these the group of the pure prefixes (de la Cruz 1973) stands out. The group includes *ā-*, *be-*, *for-*, *ge-*, *on-* and *tō-*. Instances of affixation with the pure prefixes include *abacan* 'to bake', *aberstan* 'to burst out', *abeatan* 'to beat to pieces', *acalan* 'to become frost-bitten'; *befeohtan* 'to take by fighting', *befleogan* 'to fly upon', *bebrecan* 'to break to pieces', *bebeodan* 'to offer, anounce'*;*

forberstan ‘to burst asunder’, *forbærnan* ‘to burn up’, *forbitan* ‘to bite through’, *forceorfan* ‘to cut out, down’; *gærnan* ‘to gain by running’, *gesittan* ‘to inhabit’, *gestandan* ‘to endure, last’, *gebæran* ‘to behave’, *ongēotan* ‘to infuse’, *onlūtan* ‘to bow’, *onwinnan* ‘to drink’, *ondrincan* ‘to attack’, *tōweorpan* ‘to cast down’, *tōberstan* ‘to burst apart’, *tōcīnan* ‘to split open’. It must be noted that these prefixes are remarkably opaque, as authors like Horgan (1980), Hiltunen (1983), Brinton (1986), Brinton and Traugott (2005), Martín Arista (2010) and Martín Arista and Cortés Rodríguez remark.

In this respect, Hiltunen (1983) gives an exhaustive account of the use of prefixes and pre-positional elements in Old English, which gave way to the apparition of phrasal verbs. Hiltunen (1983) distinguishes inseparable prefixes and phrasal elements. The inseparable prefixes, as the name suggests, are those that cannot appear separately from the predicate, and do not have a meaning or function on their own. On the contrary, the addition of phrasal elements to verbs gives way to the creation of phrasal verbs. It is “the analytical structure that makes it difficult to fit them into a linguistic description. In multi-word constructions, features from all of the original categories of the constituents are often brought together, and existing categories may prove insufficient to describe the new combinations” (Hiltunen 1983: 17). When a prepositional element is combined with a verb, it gives as a result a prepositional verb whose meaning can be deduced from the nature of both elements.

In Old English, the creation of phrasal verbs depends on the combination of a phrasal adverb or phrasal preposition to a verb. According to Hiltunen (1983: 20), phrasal adverbs are those “that indicate location or direction (or both) and do not normally appear as prepositions”. There are four different positions the phrasal adverb can take with respect to the verb: it can precede it or follow it, both with or without intermediate elements. Within the category *adverb*, there are, according to Hiltunen (1983) prepositional adverbs, which are those elements that function both as adverbs and prepositions, depending on the context. Hiltunen (1983) introduces this new term because “this term is more concrete than ‘particle’, and less confusing than the German-based term ‘separable prefix’” (Hiltunen 1983: 21). Compared with

phrasal adverbs, prepositional adverbs are easier to identify because prepositional functions are simpler than adverbial functions. Although the function of prepositional adverbs in the sentence is prepositional, Hiltunen regards them as adverbs, instead of prepositions, on the basis that "they involve a syntactic and, to some extent, even a semantic detachment of the prepositional adverb from the headword" (Hiltunen 1983: 22). Thus, the distinction is based more on the use and on what they imply, rather on the function they have by themselves.

As regards dependent elements, the so-called inseparable prefixes are distributed into two groups: those that have been reduced to bound morphemes, and those that may also appear independently of the verb. However, this is not a very suitable definition as the spelling of words very much depended on scribal preferences. Sometimes prefixes appeared separated from the verb, and sometimes attached to it.

A more technical approach would be to consider prefixes in Old English, those elements which were "independent words historically, forming close semantic units with verbs. At the same time they lost their accent" (Hiltunen 1983: 47). Thus, Hiltunen (1983) follows de la Cruz (1975) and states that "in his 1975 article, de la Cruz called *a-*, *be-*, *for-*, *ge-*, *of-*, *on-* and *to-* 'OE pure prefixes' on the grounds that they are "without an etymological prepositional counterpart or with an etymological prepositional counterpart which, however, differs widely in function (p. 47)" (Hiltunen 1983: 47). An account of the meanings and functions of these elements is provided:

A-: According to BT there are accented and unaccented variants of this prefix. When unaccented, it denotes negation, deterioration, or opposition. The accented variant has several meanings: out, up, end or purpose of an action… However, the BT list is not complete "but, after all, it must be borne in mind, that the various shades of meaning are innumerable, and that even in one and the same compound it often assumes different meanings" (Hiltunen 1983: 48). The combination of *a-* with a verb is very frequent, giving BT more than 600 different verbs following this pattern.

Be-: also very frequent, with some 200 different verbs listed in BT. When attached to a verb, it intensifies its meaning, or also changes an intransitive verb into a transitive one, or gives a privative meaning. The conclusion reached by BT is that "sometimes there is no perceptible variation in the sense compared with the simplex" (Hiltunen 1983: 48-49).

For-: it changes the transitivity of a verb, making it transitive from intransitive. It may also add a pejorative connotation to a transitive verb; or denote loss, deterioration, or destruction. It is difficult to state if *for-* is a prefix or a preposition, due to the grammaticalisation it has suffered.

Ge-: there is no conceivable meaning not been attached to this prefix. Lindemann (1970) rejects the prefix being (1) empty, (2) intensifying, (3) transitivising, (4) completive, or (5) perfective.

Of-: very common as well in OE, with 160 verbs listed in BT. It had the following meanings, according to BT: intensive, unfavourable force, attainment in verbs of motion, verbs of inquiring; the force of killing with verbs of striking, throwing, falling, injury with verbs denoting rest. It may also be seen as a transitivising element.

On-: it may indicate direction, 'off' 'away from the standing point', and beginning of an action. It can have the meaning 'back'. It has survived in Modern English in the negative prefix 'un-', as a merging from 'on-' with 'un-' in Middle English.

Etymological evidence is important when studying inseparable prefixes in Old English. According to the author "the difficulty lies in the fact that the elements juxtaposed rarely appear in comparable contexts in the related languages, i.e. language-specific factors must be taken into account, because they may be significant even between closely related languages" (Hiltunen 1983: 51). Meaning is also another main aspect in the study of prefixes. Where lexical meanings and secondary meanings are in doubt, prefixes are said to be intensifying or semantically empty. It is very difficult to establish a given meaning contrastively because prefixes can perform different functions, and not giving much content information at the same time.

As regards functional, syntactical, lexical and stylistic aspects of inseparable prefixes, one cannot but remark that, gradually, the Old English prefixes became less and less functional, due to aspects such as the lack of stress, and the appearance of analytical expressions in Middle English. Some authors point out "that the weakening of the prefix (*to*-) may have been influenced by possible confusion with the infinitive particle *to*" (Hiltunen 1983: 52). Syntax in inseparable prefixes is confined to transitivisation. The syntactic division prefixes had, was restricted to their composition with verbs. "In a way, the syntactic dimension of the prefixes was revived in the analytical constructions that replaced them" (Hiltunen 1983: 53).

In this respect prefix variation is relevant in the study of prefixes in OE. The fact that the same verb could appear with two or more different prefixes without much semantic variation is usually understood as a lack of expressive content in the prefixes, and the reason for their future decline. Few studies have been carried out on this respect, and they are not deeply immersed in all prefixes. Lindemann (1970: 34) points out that "because real preverbs do not express a concrete notion, but a relation that is abstract, formal and logical, they may, if they are genuine prefixes, even be substituted for one another". And continues, "this has always held generally true for the Germanic dialects, and was apparently widespread in the earlier stage of their development" (Lindemann 1970: 34). Lindemann (1970) implies that it is assumable that in a prefix-based language, interchangeability of prefixes must be common.

Other relatively frequent prefixes that attach to strong verbs to form other strong verbs are *mis*- (*miscweðan* 'to speak ill', *misdon* 'to do evil', *misfaran* 'to go wrong') and *un*- (*unbindan* 'to unbind, loosen', *undon* 'to undo', *unlucan* 'to unlock', *unwreon* 'to uncover').

Focusing on suffixation, the suffixes that are attached to form nouns include:

-dōm	*sceacdōm*	'flight'	*scacan*
-ð	*mǣð*	'cutting of grass'	*māwan*
-el	*stǣgel*	'steep'	*stīgan*

-end	*bannend*	‘summoner’	*bannan*
-ere	*crēopere*	‘cripple’	*crēopan*
-estre	*bæcestre*	‘baker’	*bacan*
-icge	*galdricge*	‘enchantress’	*(ge)galan*
-ing	*beorning*	‘incense’	*biernan*
-lāc	*brēowlāc*	‘brewing’	*brēowan*
-lic	*gelumpenlic*	‘occasional; suitable’	*limpan*
-līce	*gescēadenlīce*	‘severally’	*scēadan*
-nes	*frignes*	‘interrogation’	*frignan*
-scipe	*dearrscipe*	‘rashness, presumption’	*durran*
-ung	*fēowung*	‘rejoicing, joy’	*fēon*
-wist	*lytwist*	‘deception’	*lūtan*

The suffixes form adjectives from strong verbs include:

-fæst	*gemægfæst*	‘gluttonous’	*magan*
-ful	*sacful*	‘quarrelsome, contentious’	*sacan*
-ig	*rēotig*	‘mournful, sad, tearful’	*rēotan*
-leas	*saclēas*	‘innocent; unmolested, safe’	*sacan*
-or	*slipor*	‘slippery, filthy; unsteady, shifty’	*slīpan*

Finally, the suffixes that form weak verbs from strong verbs include:

-ettan	*bliccettan*	‘to glitter, quiver’	*blīcan*
-læcan	*gedrēoglæcan*	‘to put in order, regulate, arrange, attend to’	*(ge)drēogan*
-lian	*spearnlian*	‘to spurn, kick,	*(ge)spurnan*

		sprawl’	
-nian	*(ge)āgnian*	‘to own; claim; appropriate, usurp; make over (to); dedicate, adopt; enslave’	*āgan*
-sian	*(ge)hrēowsian*	‘to feel sorrow or penitence; do penance’	*(ge)hrēowan*

To close this section, we focus on compounding. Previous research in Old English compounds comprises works like Bergsten (1911), Borowski (1921), Carr (1939), Gardner (1968), Guarddon Anelo (2009), Talentino (1970) and Torre Alonso (2009). Kastovsky (1992) remarks that compound verbs can be of two types: derivatives from nominal compounds as in *cynehelman* ‘to crown’ and combinations with adverbs or adpositions. Strong verbs appear mainly in the second type of compounds, which can be broken down into two different categories: separable compounds (*adūnfeallan* ‘to fall down’, *forðbringan* ‘to bring forth’, *āweggān* ‘to go away’, *infaran* ‘to enter’, *onwegādrīfan* ‘to drive away’, *tōclifian* ‘to cleave to’) and inseparable compounds (*bebūgan* ‘to flow around’, *efencuman* ‘to come together, agree’ *eftcuman* ‘to come back’, *forcuman* ‘to come before’, *fullgrōwan* ‘to grow to perfection’, *oferfeohtan* ‘to conquer’, *tōberan* ‘to carry off’).

2.5. Concluding remarks

This chapter has gone from a description of the processes of word formation in general and those affecting Old English in particular, as far as the strong verb is present in the derivation. Besides, it has offered a thorough review of the problems presented in carrying out lexicological and lexicographical work upon a dead language, which moreover expands over a remarkable span of

time and is subject to different sources of variation, internally, externally and historically motivated. In this respect, the chapter has provided the reader with a description of the verbal system of Old English, with special attention to the strong verbs, that, because of their antiquity, were still very influential in the word formation of Old English. The form, function and structure of the strong verb system of Old English cannot be understood without observing its roots and its evolution throughout history. In this vein, the chapter has provided a description of the strong verb system in the different languages prior to the birth and development of Old English, namely Indo-European and Germanic. Moreover, I have put forward the different manners in which several authors have approached the study of Old English strong verbs, and the divergent classifications so far proposed. Being most of them biased by the particular focus of each work, I have opted for following the traditional classification of the strong verbs, that being a seven-class organisation compiling both *umlauting* and *reduplicating* verbs.

Once the lexicological foundations have been laid, the following chapter will introduce the concept and subject of study of lexicography. In so doing, a distinction between traditional lexicography (dictionary creation) and electronic dictionary (corpora and database compilation) will be put forward. Finally, the methodological foundations for the development of the electronic tool developed for the compilation of the Old English strong verb forms and the steps taking in conducting this research on their lemmatisation will be provided.

3. LEMMATISING OLD ENGLISH STRONG VERBS IN A LEXICAL DATABASE

3.1. Introduction

While chapter 2 has presented the subject of lexicology as a branch of linguistics and discussed some linguistic properties of Old English relevant to this work, the current chapter sets the basis for the lexicographical part of the undertaking. This is tantamount to saying that once the theoretical aspects of the research have been introduced in the previous chapter, this one turns to the applied side of the question. Thus, the progressive evolution of lexicographical work towards electronic lexicography is discussed in section 3.2 in order to provide this dissertation with a coherent and, furthermore, up-to-date framework. Section 3.3 reviews the most outstanding works in Old English lexicography, including *An Anglo-Saxon Dictionary* (Bosworth and Toller 1973), *The student's Dictionary of Anglo-Saxon* (Sweet 1976), *A Concise Anglo-Saxon Dictionary* (Clark Hall 1996), and *The Dictionary of Old English in Electronic Form A-G* (Healey et al. 2008). Against this background, section 3.4 presents the relevance, aims and scope of this research. Then, section 3.5 discusses the methodological decisions and steps taken to carry out the research. Section 3.6 engages in the discussion of the problems and solutions put forward for different aspects of the research. Finally, some concluding remarks are made in section 3.7.

3.2. Traditional lexicography and the growth of e-lexicography

As has been pointed out in chapter 2, lexicology focuses on the analysis of the morphological, functional and semantic properties of words and the word formation processes. Lexicography, on its part, is an empirical science that makes use of the analysis and explanations lexicology provides in order to classify the words of a language and organise them in dictionaries or databases. Its main field of study is

the inclusion of words into dictionaries, thus paying special attention to the correspondence of form and meaning. Dictionaries arise as tools that allow speakers to consult the meanings and orthography of words in their own language, in the case of monolingual dictionaries, and to look up words from other languages, in bilingual dictionaries.

The core use of the term *lexicography* makes reference to the elaboration of dictionaries, and to the study of dictionaries and their use. Although mostly based on language, there are several types of dictionaries with different focus as stated by Coleman (2006: 581), who draws attention to the dictionaries of "human rights, the Middle Ages, African-American architects, nuclear engineering, dreams, birds, love, and medical quotations, among others". These types of dictionaries come from an Internet search for publications from 2003 with the word *dictionary* in their title. According to the author, this indicates that the term *dictionary* suggests that the information is arranged alphabetically, and includes explanation and exhaustiveness, something that is closely related to the term *encyclopaedia*.

Bearing this in mind, researchers have tried to arrange dictionaries considering their structure, content, use, size, information categories included, and languages covered, among other aspects. For Coleman, *dictionary* means "a reference work dealing with words" (Coleman 2006: 582). This definition includes thesauri, but excludes encyclopaedias. In this category, proverbs, idioms, etymology, new words, and personal and place names can be included. Another possible categorisation may include the intended uses of those dictionaries. There are also "college dictionaries, school dictionaries, beginners', elementary, intermediate and advanced dictionaries, illustrated dictionaries, and picture dictionaries" (Coleman 2006: 582). Some dictionaries focus on a particular use, not on the type of user, as there are rhyming dictionaries, crossword dictionaries… Others are restricted to the type of vocabulary they include, and in this group it can be found dictionaries of slang and euphemisms, dictionaries of synonyms and antonyms, dictionaries of British, American, Australian (etc.) English. In addition to all these possible categorisations of dictionaries, there is another classification based on size, and within this group there are desk, pocket, mini, super-mini, concise, compact, or unabridged dictionaries.

Dictionaries are not only research tools, but also subjects of study. When speaking of dictionary research, most lexicographers have focused their effort on dictionaries published before 1830, the *Oxford English Dictionary* and modern monolingual learners' dictionaries.

An important area of dictionary research was the history of the English dictionary. In this respect, Starnes and Noyes's (1946) study of the early history of the English dictionary is unrivalled, but it has been supplemented, mainly by Hüllen's (1999) account of early thesauri, Cowie's (1999) examination of learner's dictionaries, and Gotti's (1999) and Coleman's (2004a, 2004b) studies of slang and cant dictionaries. Studies of individual dictionaries include Reddick (1996) on Johnson's dictionary, Micklethwait (2000) and Morton (1994), on Noah Webster's *American Dictionary of the English Language*, and Mugglestone (2000) and Willinsky (1994), on the *Oxford English Dictionary*.

Researchers are broadening the scope of their field chronologically and textually. The study of manuscript dictionaries has challenged the existing preconceptions of the English dictionary. Before this, the focus has been placed on later printed works. Historical dictionary studies try to understand the relationships between early monolingual and bilingual dictionaries.

Learner's dictionaries include monolingual and bilingual dictionaries. Bilingual dictionaries have a longer history, compared to monolingual, dating back to the Old English period. However, monolingual dictionaries have been proved to be more useful to users, as they provide more detailed information about the language to be learned. There has been a growing interest in the history of modern learners' dictionaries. Cowie (1999) points out three generations of monolingual learners' dictionaries. The first generation was influenced by the vocabulary control movement which defined English words using a restricted range of vocabulary; the second generation is interested in phraseology; and the third is centred on the use of computer technology. The major learners' dictionaries are under constant revision, and details can be found at their websites.

The user perspective is an important focus in recent dictionary research. Lexicographers improve dictionaries by understanding the actual use of dictionaries. Fraser (1997) says that "in dictionaries aimed at native-speakers, pronunciation

guides are most usefully given in the form of non-phonemic respellings […] rather than in the more commonly used International Phonetics Alphabet (IPA)". Many studies have focused on learners' dictionaries, and how and why learners of English consult dictionaries, and how the dictionaries need to be designed to allow the users to locate the information contained in the dictionaries.

Advances in information technology improved the development of learners' dictionaries. According to Coleman (2006: 584), "Ooi (1992: ch2) makes a useful distinction between computational linguistics (the building of lexicons for natural language processing), computational lexicography (the production or use of machine-readable dictionaries), and computer corpus linguistics (the principles and practice of compiling bodies of electronic texts of actual language)". The use of language corpora, instead of the creation of a word list, has brought about a revolution in the production of dictionaries. The current trend is to collect texts and recordings of English in use to reflect the actual usage rather than lexicographers' preconceptions or linguistic history.

Dictionaries are, and have been used, as cultural products and political tools. Dictionaries show their cultural biases "in their selection of headwords, usage labels, and citations, in the wording of definitions and ordering of senses, and even in their willingness to give etymologies from particular language sources" (Coleman 2006: 584). Examples of dictionaries serving a political purpose are Thomas Spence's *Grand Repository of the English Language* (1775), which proposed a reformed alphabet to enable lower classes to achieve sufficient literacy to become politically aware; and more recently, the production of Canadian English dictionaries, which, according to Lilles (2000), have played an important role in the development of national and linguistic identity.

Internet has favoured the publication of online dictionaries. Dictionaries have no longer a fixed form, and the edition of entries is easier and more immediate. Some online dictionaries do not edit the submitted material by their users. Some sites allow the consultation of several online dictionaries simultaneously. The use of online dictionaries is more flexible, as entries can be reordered or searched according to the user's needs: understanding a document, struggling to complete a

crossword, or attempting to write a poem. Internet has also helped to facilitate the publication of research materials that would not be economically viable in other way.

Lexicography as a discipline has been professionalised and there are training programmes and degree courses available for lexicographers. Hartmann (2001: 7) notes that "in the past people drifted into and out of dictionary making from such diverse occupations as theology and education… literature... philology… medicine… and music…". Nowadays, MA degrees in lexicography are available at many universities, and many undergraduate linguistics programmes include courses on lexicography as well. Publications based on lexicography include manuals for lexicographers, accounts of developments in lexicographic practice, and textbooks for teachers, researchers, and students. However, these practices are not free of controversy. Authors like Gusun (1998) discuss that lexicography is a discipline on its right and he goes on to discuss whether it is advisable that lexicographers theorise about their own work or not.

But dictionaries are not only analysed for their subject of study or for the goal they are designed to fulfil. Dictionary content is also a crucial topic under study in lexicography. Coleman (2006) compares the entries to different dictionaries for the word *dictionary*. The differences between dictionaries range from length, to the features dictionary editors take from previous dictionaries, through the presence or absence of etymological information, or the addition of semantically related terms, among other aspects. Some dictionaries, as the *OED*, offer examples of compounds and derivatives. Citations only illustrate the use of the term, rather than telling us anything about its history.

Senses in lexicography are based on the perceived needs of the dictionary's intended users and on practical constraints such as the projected size and cost of the dictionary. Most modern dictionaries recognise the bilingual dictionary as a different type of reference work, as it provides the sense with a separate numbered definition. The *Oxford English Dictionary* lists several senses that none of the other dictionaries include. This dictionary, as it is historical, orders its senses chronologically. Learners' dictionaries order senses by frequency, so the most commonly used sense of a word will come first.

In addition to the study of the senses a dictionary can include, it is important to study as well the structure of the dictionary entry, breaking it down into its constituent elements. Dictionaries provide a wide range of different types of information about the listed words, the only feature shared being the headwords themselves. Neither phonetic nor grammatical or etymological information are essential to dictionaries, although the dictionaries that do not indicate part of speech imply it in their definitions. Etymologies are not included in most of the modern works, particularly those aimed at learners or providing restricted coverage. The fact that definitions are not essential either is somehow surprising. Coleman (2006: 591) points out the fact that some dictionaries provide indirect definitions by means of related words, thus "Harraps and Partridge instead provide a synonym in French or standard English. Phillips joins them in this, and […] gives us a synonym in Greek". Finally, the use of dates to indicate the inclusion of a word in the lexicon of a language is only proper of historical dictionaries and is not regulated. Coleman underlines this point by reviewing two different works and shows that "Partridge gives only an indication of period, while the *OED* provides specific dates" (2006: 591).

Apart from their academic interest, most dictionaries are commercial products, and need to have market demand. Problems may arise if a particular product does not match its expected demands. In this sense, (Stein 2002: 34-35) remarks the failure of the third edition of Webster's dictionary of American English as "the editors tried to reflect the way English is used […] and the buyers expected to be told what was right and wrong – and more to the point, they did not want to see usages that they regarded as wrong legitimized by inclusion."

All in all, lexicography and dictionary research are undergoing major changes, most of them thanks to the development of information technology. Computers and the Internet offer endless possibilities. The market of dictionaries is still growing, and it is becoming more specialised, as researchers review all these works and study how they can feed back future generations of dictionaries.

In this vein, as corpora keep growing, and the work of lexicographers is becoming so thorough that they require new computational technologies to make their work accessible and profitable, electronic lexicography plays an essential role

in the creation of new dictionaries. The trends now focus on the creation of online dictionaries, and CD-ROM dictionaries, rather than printed dictionaries. This situation has given the opportunity to lexicographers, along with computer engineers, to create new ways of interacting with the dictionary, so new methods of corpus encoding, such as search strings, have come up to help lexicographers to access to the data, making it easier to search and classify all the data filed in the database. This preference for the digital dictionary is encouraged by the slow decline of the use of printed (physical) material, due to the increase of space resulting from the extension of the language in terms of new forms, and a widening in the meanings words have. So this increase in the number of forms has triggered the creation and development of new devices capable of filing and dealing with the language(s) under study, something that electronic and online supports can achieve. New software and procedures have been developed to fulfil the new demands electronic lexicography began to request. This combination of traditional lexicography and electronic device development has modified to a certain extent the work of the lexicographer, who must understand how database search engines work, and need to know how to code the dictionary entries so that the search engine can process them and turn out results depending on the search string the user may enter to look up a certain entry.

Apart from the query question, Granger and Paquot's (2012) compile a series of articles on the new trends and directions of electronic lexicography. They acknowledge the process through which computer technology has been applied into lexicography and dictionary creation. According to Granger and Paquot (2012: 15-16), "for some time, lexicographers have been struggling with the constraints of print: with access to powerful corpus-querying software applied to billion-word corpora, we have the tools (and the data) to provide a fuller and more systematic account of how language works".

The first aim of lexicographers working with computers is the creation of a corpus large enough to fit their needs in order to offer an approach to the language as exhaustive as possible. Compared to earlier days, it is easier to work with larger corpora thanks to computers. The drastic reduction in the cost of electronic devices has strongly contributed to the fast growth of the discipline. In Granger and Paquot's

(2012: 18-19) words, “quite suddenly, a number of factors combined to make it possible, at relatively low cost, to collect, annotate, and store corpora measured in billions of words rather than millions”. In fact, the reduction of economic costs has played an important role in the democratisation of dictionary creation, by allowing lexicographers to have heavy investments from companies, and by allowing users to access dictionaries at an affordable cost. The focus has moved from the aim of publishing dictionaries to user accessibility.

On the technological side, as the corpora grew, so did the needs of the lexicographers, who in order to deal with larger corpora in a more efficient way, ended up with the creation of more intelligent corpus-analysis software, which helped them to analyse aspects such as word collocations, in a faster and easier way. According to Granger and Paquot (2012: 22) “not only did this bring welcome efficiencies in the dictionary-making process, the continued growth in corpora became an advantage instead of a problem: higher volumes of data helped to make the sketches an ever-more reliable reflection of real usage, but without adding to the lexicographer’s workload”. These authors assume that in the future machines will to a certain extent replace lexicographers, provided that technology keeps developing at the same rate. Nonetheless, Granger and Paquot (2012) still acknowledge the need for a trained specialist when stating that “the hardest parts of lexicography - word sense disambiguation, definition-writing (for monolinguals), and providing translation equivalents (for bilinguals) - still require expert intervention by skilled lexicographers” (Granger and Paquot 2012: 29). They conclude that the combination machine-lexicographer will remain unaltered for many years. Computers are essential in this new way of storing language and creation of dictionaries, but a human brain is still needed to properly understand language, something that cannot be done by computers in their present state of development.

Focusing on the corpora compilation and on the electronic corpus tools, Kilgarriff and Kosem (2012: 31) state that “compiling and storing corpora has become faster and easier, so corpora tend to be much larger than previous ones”. These electronic tools are faster, more multifunctional and customisable. For Kilgarriff and Kosem (2012: 31) “corpus tools had to be improved to assist lexicographers in adapting to this change”. Electronic lexicography very much

depends on automatic analysis. Thus, the corpora have to be prepared. This preparation is done in two steps: "preparing the *metadata*, or *headers*, and preparing the *text*" (Kilgarriff and Kosem 2012: 32). The main work of the lexicographer in this step is to examine the instances of a word if they want to know "from what kind of text a particular instance has been extracted" (2012: 32). *Headers* are useful at this point, as they give useful information about the features of the document in such a way that the corpus tool is capable of interpreting them. With these *headers*, corpus tools can provide information on the texts, and also limit the searches to particular text types, create word lists and find keywords, among other tasks. When preparing the text, the lexicographer needs to start by managing the character encoding, and then marking the texts with "(1) sections, paragraphs and sentences, (2) tokens, (3) lemmas, (4) part-of-speech tags, and (5) grammatical structure" (Kilgarriff and Kosem 2012: 32). Character encoding is helpful in order to mark which elements are language-specific or writing-system specific, but this encoding may create problems of "misinterpretation where the system assumes that one encoding has been used, but in fact a different one was involved" (Kilgarriff and Kosem 2012: 32) as each language requires different encoding system.

Corpus tools may be categorised according to a series of features. Kilgarriff and Kosem (2012) propose the following subdivisions:

- Computer-based (stand-alone) tools vs. online tools: the former tools require that "the tool and the corpus are stored on the user's computer" (Kilgarriff and Kosem 2012: 33). Computer-based tools are WordSmith Tools and MonoConc Pro.
- Corpus-related tools vs. corpus-independent tools: "some corpus tools can be only used with a particular corpus, most often because they were designed as a part of a specific corpus project or for a specific institution" (Kilgarriff and Kosem 2012: 33). Corpus-independent tools can be used to analyse any corpus.
- Prepared corpus vs. web as corpus: most of "corpus tools are used to access a corpus that has been compiled with linguistic research in mind. But the web can be viewed as a vast corpus, with very large quantities of texts for many

languages, and lexicographers frequently use it in this way" (Kilgarriff and Kosem 2012: 33).

- Simple tools vs. advanced tools: as corpora are growing in size, and the number of users is increasing, "corpus tools have become more and more multifunctional, they have started offering many different features to assist their users with analysis" (2012: 34). The features of these tools range from concordance, collocation and keywords to word sketches and string search.

This classification reflects the different goals of lexicographers. On the one hand, this classification points to the different aims lexicographers may have when making a dictionary. The use of always-growing corpora needed the development of new software capable of processing such amount of information in a short period of time, so that lexicographer work was faster and easier. The generalised use of the Internet as a communicative tool is in itself a huge corpus that lexicographers frequently use to consult and withdraw information from, due to the amount of written language there found, thanks to the contributions of users and media who communicate in this way. On the other hand, the classification given above evinces the kind of user to which the lexicographical product is addressed. There are three major types of users: lexicographers, linguistics researchers and students, and language teachers and learners. For this reason, different tools have been designed depending on the characteristics of the users.

Along with the ability to store bigger amounts of information and providing more systematic and faster methods of analysis, the corpus tools provide electronic lexicography with a series of advantages with respect to traditional lexicography. As Hanks (2012: 57) puts it, "electronic lexicography opens up all sorts of radical possibilities that were closed to traditional lexicography: new kinds of evidence, new modes of description, new ways of organizing evidence, new possibilities of exploiting database structure and hypertext links, and the need for new theoretical foundations". Apart from these advantages, Hanks (2012) points out the possibility of showing the word in context, and that the spatial constraints printed the book has disappeared with the burst of the Internet. In Hanks's (2012: 57) words "lexicography is in transition; publishers, nervous about future commercial

prospects, are wary of investing in large-scale innovations, just at a time when such innovations are most needed. At the same time, funding agencies and their advisers have not yet been convinced that major innovations such as FrameNet and Corpus Pattern Analysis would justify the large-scale research investments that would be needed to bring them to completion and yield practical benefits".

When comparing electronic lexicography to traditional lexicography, Hanks (2012) points out four traditions of lexicography, namely:

- Scholarly dictionaries on historical principles, which consisted on compiling words of a certain language just to "trace their origins and semantic development" (Hanks 2012: 58). From the 18th Century up to very recently, lexicographers took it for granted that studying the origins meant studying meaning. This means that "if a word changes its meaning over time, it was assumed that the older meaning was somehow more correct than any more recent development" (Hanks 2012: 58). From the point of view of the author, "roots are culturally important [...] however, knowing about roots is not the same as knowing about meaning" (Hanks 2012: 59).
- Dictionaries of current usage for native speakers: at first, dictionaries were thought to give historical information about the origins and development of a word, but not to give the meaning of that word. This changed at the end of the 19th Century with Funk and Wagnall's Standard Dictionary of the English Language (1894-97), which "deliberately reject[s] historical principles and instead record[s] the current meaning of words" (Hanks 2012: 59). One of the main problems that lexicographers at that time faced was that when there was not enough evidence for the meaning of a word, they did not know which meaning put first as the most common. Even when the first corpora appeared, they found out that what they thought as speakers would be the most common usage of a word was not such. "Social salience (i.e. the most frequent senses of a word) and cognitive salience (memorability, or rather ease of recall) are independent variables" (Hanks 2012: 59).

- Bilingual dictionaries: bilingual lexicographers “have attempted to give practical implementation to the long-recognized fact that aiming at literal word-for-word translation between languages is a naïve goal that leads to errors - often, ludicrous errors” (Hanks 2012: 60). Phraseology started to have a more important role. Atkins and Duval started to compile ‘frameworks’ consisting of typical phraseology for each word in each language. The problem was that the equivalents of each language were not taken from any source, but created by lexicographers.
- Dictionaries for foreign learners. Horner’s Idiomatic and Syntactic English Dictionary was pioneer in creating “a work that would help learners to use the syntactic patterns and idiomatic phraseology of English with reasonable accuracy when writing and speaking” (Hanks 2012: 60). It was useful to help language production rather than having a ‘decoding’ purpose. Later on this purpose was ignored and many words were added to following editions.

In Hanks’s (2012) opinion, the appearance of electronic corpora and their application on lexicography has had a major impact upon dictionaries for foreign learners, the first of them being the COBUILD. Up to this point, new dictionaries were corpus based, but not corpus driven like the COBUILD. For monolingual lexicographers, “large corpora provide [...] with sufficient evidence to decide what to include and (more importantly) what to leave out” (Hanks 2012: 63). Corpora also help to improve the accuracy of explanations and the pragmatic uses of words and phrases, neglected in traditional dictionaries. The future of lexicography and dictionaries hinges on the use of hypertext structures and links, not on printed books.

Considering the future of electronic dictionaries and corpora, there is still a wide area of work and improvement. Able (2012) describes the general trends on dictionary writing systems. She states that although dictionaries and dictionary entries have a highly structured pattern with many common and recurring elements, there are some aspects that need to be defined: “the principles of data collection and selection, the treatment of different types of information and parts of speech, the use

of a specific metalanguage, the treatment of macro-, medio- and microstructural elements" (Able 2012: 84), among others.

A detailed style guide is useful in dictionary entries to allow dictionary users to quickly and easily find the information they are looking for. According to Atkins and Rundell (2008: 212) "computerization has largely relieved lexicographers of the need to pay attention to such tasks". This is the main task of dictionary writing systems. According to Kilgarriff (2006: 7) "a dictionary writing system is a piece of software for writing and producing a dictionary. It might include an editor, a database, a Web interface and various management tools […]. It operates with a dictionary grammar, which specifies the structure of the dictionary" However, for Atkins and Rundell (2008: 103) it is "a program that enables lexicographers to compile and edit dictionary text, as well as facilitating project management and (later in the process) typesetting and output to printed or electronic media".

Following Atkins and Rundell (2008: 114), the function of a dictionary writing system is to edit entries, and to "cope with the particular demands of complex dictionary writing projects". Three aspects acquire importance in dictionary writing. The first one is the content of the dictionary. The second one is that every dictionary should have a specific configuration of its components, or a structured data model. And the third one, the formatting and style of the data presentation. Data input happens, in principle, by entering the text in one sequence into a word-processing system. By following this system, font type and size are selected directly, among other aspects, and the final dictionary text can be easily processed and stored, as it may appear in the final product. This allows flexibility to data input. Among the disadvantages of the writing systems, Able (2012: 88) points out that "different types of information are not always explicitly marked and thus the final product is not searchable; automatic checking of consistency and conformity can also only be done in a limited yaw. In addition, the reusability of data is not straightforward".

Recent developments have made a fundamental change in lexicographic processes, which are more automatised and now lexicographers compile databases and update their existing material. The dictionary writing system has now become central in the production of dictionaries, as it is increasingly versatile, multifunctional, and with processes that can be controlled and easily implemented.

In this respect, it is important to create an understandable code for lexicographers when trying to enter information in the database from which the dictionary is created. It aids lexicographers to enter information quickly and easily, more than in the old way of processing information, and it is a useful tool to standardise all the procedures and codes used for the purpose of compiling information. In this sense, Svensén (2009) and Able (2012) agree in signalling XML mark-up language as one of the best writing systems. Using XML mark-up language helps lexicographers to input, organise and edit the data. Although the data is organised hierarchically, XML allows data to be stored as a file and as a database, and it "ensures quicker and more direct access to the data than other storage systems" (Svensén 2009: 421). As explained above, using a writing system is not only helpful for the creation of the database, but also to update and improve all the information already existing in the database, so all the possible changes that the database may undergo once it is created are made in a more efficient way than before.

As has been remarked above, the change from traditional lexicography/printed dictionaries to electronic lexicography/corpora is still on the making, and there are many challenges yet to be faced and overcome, especially on the theoretical side. According to Tarp (2012: 107) "printed dictionaries will be published for a long period ahead but, at the same time, it is no secret that the electronic medium is gaining still more ground and will gradually overtake and outshine paper as the preferred platform". Tarp supports the idea that a new lexicographical theory based on the use of e-tools is needed, although there is some controversy on the topic. Tono (2012: 2) wonders whether "we really need a 'theory'" although he finally answers affirmatively to his own question.

After reviewing the theory underlying lexicography, Tarp (2012) states that there are general and specific theories. Among general theories, the functional theory stands out from the rest, although, following Sorokoletov (1978), Tarp (2012) acknowledges the importance of the Russian school. In Tarp's words (2012: 108-109) "in the Soviet period lexicography developed into an independent discipline with its own theory, own tasks and own methods for their solution". On the other hand, specific theories reflect topics and sub-areas related to the new electronic platform and the new options available for lexicography.

For Tarp (2012), lexicography is a discipline that has been present in our society for about four thousand years. The only thing that changed is the medium; we have gone from clay, papyrus or paper to the Internet, CD-ROMs and mobile phones. According to Tarp (2012: 109-110), "if a general theory of lexicography in the real sense of the word had existed four thousand years ago, there would have been no reason to change it when the practical tools of lexicography passed from clay to papyrus, and later from handwritten to printed […] the only thing theoretically new to be developed would be […] specific theories related to the new media, for instance specific theories about data processing, data presentation, data access, and data linking". As it comes out, any theory needs long periods of time to be developed and/or even changed for a new one, as is the case of lexicography, in which the research field changes through time, so the discipline has to adapt itself together with the field of study.

Regarding the implications of electronic lexicography for the creation of dictionaries, Tarp (2012: 113) states that electronic dictionaries constitute "new forms of access and presentation of the selected data". When the question of *freedom from space constraints* arises, Tarp refers to the *Yongle Dadian* and the *Gujin Tushu Jicheng*, two Chinese works that are so gigantic that seem to ignore the space constraints. In this regard, Lew (fc.) offers a distinction between *storage space* and *presentation space*: In his words, "on careful inspection, it appears that the notion of *dictionary space* is not specific enough as a technical term because it is ambiguous. The suggestion that dictionary space is unrestricted is actually largely correct, but only when space is understood as the capacity to hold the total content of the dictionary - this sense of *dictionary space* could provisionally be called *storage space* […] *presentation space* refers to how much can be presented (displayed, visualized) at a given time to the dictionary user" Lew (fc.). According to the author, a new adaptation of lexicography to the new possibilities that electronic media and information science offers is required.

Closely related to the presence of e-tools in lexicography is the individualisation of the discipline. This individualisation stands for the new individual needs that every user may have when consulting an electronic dictionary, "this is the reason why *the individualization of user-needs satisfaction* is a question

to be taken seriously at a time when computer and information sciences are gradually providing the necessary technology that may allow this gigantic and revolutionary step in the framework of renewed lexicographical theory and practice" (Tarp 2012: 115).

Dictionary creation in the traditional way needs a huge space to have all the language compiled, and this space may present problems when the user faces it. Electronic dictionaries in this respect are also important as the space is virtualised, and it stops to be a problem for users and also for lexicographers and publishers. It is easier for lexicographers to deal with huge corpora that do not occupy a given physical space, where the access to the information is simpler. For users, they have more simplicity to consult the information. And finally, publishers find it cheaper to invest and create dictionaries, as they do not have to spend on physical supports as paper, so dictionary creation is more and more affordable when the dictionary is presented electronically.

By way of conclusion, this section has made the point that the field of lexicography is undergoing a great transformation due to the incorporation to lexicographical work of technological devices and electronic corpora. As a result of this transformation, the access to information has been simplified and the information given by dictionaries has been extended. There has also been a process of specification and specialisation of the purpose dictionaries have, including morphological and syntactic information, together with the use of corpora to compile and retrieve information. The use of corpora is practically generalised in dictionary creation, as it guarantees the use of exhaustive and accurate information. For the management of large collections of linguistic data, new software has appeared to help the work lexicographers used to carry out previously, so this work seemed to have been reduced and passed to computers. Bearing this in mind, together with the work of lexicographers, the work of computer engineers is playing nowadays a crucial role, since they provide the database software. It can be said that the dictionary creation process is a multitask one, including lexicographic work with computing processes, which combined together result in an electronic dictionary that can be easily accessed by users and professional lexicographers in the same way,

can be expanded, modified or adapted to new goals or needs and, moreover, is accessible on the Internet.

3.3. Review of Old English dictionaries

Although the present and future of lexicographical studies are based on the application of the new technologies we must acknowledge the importance of the traditional methods of investigation and the wealth of philological data thus compiled. This section provides a review of some of the lexicographical sources available for the study of Old English. The review of these works not only provides a basis for the current study but also reinforces the relevance of this undertaking. On the one hand, it allows me to get into the details of lexicography on the subject of study. On the other, it opens the gate to the observation of similarities and differences in the data, classification and distribution as presented by the different authors in their dictionaries. Finally, the study of these sources allows me to compile an initial list of strong verbs which will constitute the starting point of the search for inflectional forms.

This section reviews four of the main lexicographical sources of Old English currently available, including *An Anglo-Saxon Dictionary* (Bosworth and Toller 1973), *The student's Dictionary of Anglo-Saxon* (Sweet 1976), *A Concise Anglo-Saxon Dictionary* (Clark Hall 1996), and *The Dictionary of Old English in Electronic Form A-G* (Healey *et al*. 2008). Ellis (1993), in spite of pointing to some inconsistencies and shortcomings of *An Anglo-Saxon Dictionary*, states that "the Bosworth-Toller dictionary is far superior to Bosworth´s earlier work, and together with Toller´s 1921 *Supplement*, this work remains the most comprehensive Old English dictionary currently available".

The four dictionaries cited above share the fundamental characteristic of trying to present "headword spellings as they are most commonly found in Old English texts" (Ellis 1993: 5). In practice, this means that they are more focused on the West Saxon variety of Old English than on the other varieties, as grammars in general do (thus Campbell 1987; Quirk and Wrenn 1994; Hogg 2011; Hogg and

Fulk 2011). This is a consequence of the dialect of the written records, which is mainly the West Saxon one. Also due to the availability of written evidence is the fact that *The Dictionary of Old English in Electronic Form A-G* (henceforth DOE) represents the spelling of late texts, most of which are written in West Saxon. Although these dictionaries are geared to West Saxon, they also account for the records written in the other dialects.

Turning to the differences between the dictionaries at stake, they differ in format, goals, scope, organisation and degree of exhaustivity. All dictionaries, except *The Dictionary of Old English in Electronic Form A-G*, which is accessible and searchable online, have been published in paper. An online version of *An Anglo-Saxon Dictionary* (hereafter Bosworth-Toller) is available at http://bosworth.ff.cuni.cz but its functionalities do not compare to those of the DOE. There are also digitised versions of *The student's Dictionary of Anglo-Saxon* (hereafter Sweet) and *A Concise Anglo-Saxon Dictionary* (henceforth Clark Hall), although they do not make any difference with respect to the paper version. All dictionaries are comprehensive or unabridged, except Sweet, which has been reduced in size by means of the omission of terms and definitions. The preface to Sweet makes it clear that the dictionary was born out of the request by the Delegates of Clarendon Press to publish an abridgement of Bosworth-Toller. Henry Sweet (1976: vii) justifies the decision on practical grounds by remarking that "if [a dictionary-DMR] is done ideally well and on an adequate scale it is never finished – and an unfinished dictionary is worse than useless- or, if finished, is never uniform as regards materials and treatment". Leaving aside the scope, Sweet has a clear advantage over the others in the fact that it is the only one that arranges entries not only alphabetically but also by word family. All dictionaries reviewed in this section are complete, with the exception of the DOE which, as its title indicates, has reached the letter G. With the exception of Bosworth-Toller, which does not always lemmatise, thus including numerous inflected forms as headwords (typically past participles or irregular forms), all dictionaries lemmatise thus unifying all inflectional forms under the corresponding headword. It is worth noting in this respect that the DOE includes the infinitive and the past participle of verbs on a regular basis. Numerous differences arise that are related to alternative spellings.

Sweet contains fewer spelling alternants, and fewer inconsistencies and circularities in this respect. On Bosworth-Toller, Ellis (1993: 4-5) comments:

> While the Bosworth-Toller dictionary is also, without a doubt, more systernatic than Bosworth's earlier work, it still suffers from some of the inconsistencies in spelling and arrangement of headwords found in Bosworth´s *Compendious Dictionary*, particularly in the treatment of orthographic variants and in a consistent method of cross-referencing (...) following the *vide* back to the main entry sometimes can lead the reader on a frustatingly circuitous route. For example, at *ciele* ('cold') the reader is directed to the alternate spelling *cile*; at *cile* the reader is directed to his final destination, *cyle*.

These aspects are discussed in more detail in the sections that follow: headword spelling (3.3.1), scope (3.3.2), alternative spellings and cross-references (3.3.3), phonology (3.3.4), syntax (3.3.5), morphology (3.3.6), meaning definition (3.3.7), etymology (3.3.8) and textual evidence (3.3.9).

3.3.1. Headword spelling

Perhaps the first problem addressed by a lexicographer when facing the creation of a dictionary of Old English is determining the headword spellings that are going to define the dictionary entries. In a language where a variety of spellings are available, this task becomes crucial, for it defines the first and foremost property of the dictionary. Ellis (1993) addresses this problem, and states that, although different techniques, methods and approaches have been used so far by different authors to face this problem, since the publication of *The Students' Dictionary of Anglo-Saxon* (Sweet 1976) a general tendency to turn the Early West Saxon dialect into the standard variety of Old English has been developed. Ellis (1993) in his review of the problems of Old English headword spelling follows Wrenn (1933: 82) in acknowledging the usefulness of Sweet's normalisation for teaching purposes. However, Sweet's (1973) system is not free of problems. Despite his attempt to obtain an idealised, normalised standard of Old English, based on the Early West Saxon dialect, problems arise in different fronts. On the one hand, only three texts

from the Alfredian period (late 9th-early 10th centuries) are available. On the other, the lack of diatopic and diachronic influence put assessed by Sweet, can be argued. Wrenn (1933) already acknowledged the unsystematicity of his model. In this vein, Ellis (1993) summarise the problem in the following table:

Sweet (1871)	Sweet (1976)	Late West Saxon
all	eall	eall
haldan	healdan	healdan
bion	beon	beon
biorht	beorht	beorht
monig	manig	manig
monn	mann	mann
biscep	biscep	bisceop
hefon	heofon	heofon
burg	burg	burh
sorg	sorg	sorh
fierd	fierd	fyrd
hiene	hiene	hine

Table 1: Headword spelling variation in *The Students' Dictionary of Anglo-Saxon* (from Ellis 1993).

Table 1 shows the inconsistencies in the headword spelling system of Sweet's (1976) dictionary. The table compares the spellings found in the dictionary against the standards Sweet himself proposed in this *Pastoral Care* (1871) and the spelling of Late West Saxon. As seen in the table, Early West Saxon spelling is maintained in Sweet (1976) in a limited number of words. According to Ellis (1993), the words listed above are the only words in Sweet (1993) retaining spelling properties of that period. Those properties are the use of the unpalatalised <g> and the diagraph <ie>,

which is, in fact, the only property that can be exclusively attributed to the West Saxon dialect. The first half of the table reflects the fact that, despite his initial purpose, Sweet (1976) makes use of spellings which are proper of a later period.

The importance given by Sweet to Early West Saxon (1976) leads to what Wrenn (1933: 67) named "mechanical oversystematizing". This concept implies the lack of alternative spellings and cross-references, as well as the use of headwords spellings which are not assessed in any word form. The lack of cross references leads Sweet to present the words *ceald* and *cield* both meaning "cold" as separate words rather as variant spellings of the same word.

As regards his preference for the use of the <ie> spelling, even if unattested, turned into an artificial system which was followed, and increased by other authors. According to Ellis (1993), Holthausen's (1963) etymological dictionary, adopts Sweet's (1976) <ie> spelling and even goes beyond, using that spelling in words in which Sweet (1976) makes use of a more modern spelling, as it is the case of *ciecen* "chicken", which Sweet lists as *cycen*.

Other dictionaries, such as Clark Hall (1996), also attempt to use Early West Saxon spelling, but it is more unlikely to include unattested spellings. On the opposite extreme is the DOE, which follows a totally different path, and prefers to choose the oldest attested form for its headwords.

As way of summary, Ellis (1993) provides a comparison between the different spellings of some headwords in different dictionaries, namely Bosworth (1848) *A compendious Anglo-Saxon and English Dictionary*, Bosworth-Toller (1973), Sweet (1976), Clark Hall (1996) and Cameron, Amos, Healy *et al*. (1986-91).

	Bosworth	**Bosw-Toller**	**Sweet**	**Clark Hall**	**DOE B-D**
to beckon	bicnian/ bycnan	beacnian	biecnan	biecnan	bicnan
to bend	bigan	bugan	biegan	biegan	bigan
arch, vault	bigels	bigels	*biegels	bigels	bigels
to encourage	byldan	byldan	bieldan	bieldan	byldan

boldness	byldo	byld(o)	bieldo	bieldo	byldu
a trumpet	byme	byme	bieme	bieme	byme
birch	birce	birce	*bierce	beorc	byrc
to shine	---	beorhtan	bierhtan	bierhtan	beorhtian
brightness	bierhte	beorhtu	bierhto	bierhtu	beorhtu
to burn	byrnan	beornan	biernan	beirnan	byrnan
mallet	bytl	bytl	bietel	bietl	bytel
concubine	cyfes	cyfes	ciefes	cifes	cifes
to call	cygan	cigan	ciegan	ciegan	cigan
cold	cyle/cile	cyle	ciele	ciele	cyle
merchant	cypa	cypa	ciepa	ciepa	cypa
onion	cipe	cipe	ciepe	cipe	cipe
a shout	cyrm	cirm	*cierm	cirm	cirm
to turn	cirran	cyrran	cierran	cierran	cyrran
fastidious	cis	cies	*cies	cis	cies
cheese	cyse	cyse	*ciese	cyse	cyse
coffin	cist	cyst	*ciest	cist	cist
kettle	cytel/cetel	cytel	*cietel	citel	cytel
to kill	dydan	dydan	*diedan	dydan	dydan
to dip	dufan	dufan	*diefan	dyfan	dyfan
hidden	diogol	digol	diegle	diegol	digol
to dip	dyppan	dyppan	*diepan	dyppan	dyppan
to conceal	dyrnan	dyrnan	diernan	diernan	dyrnan

Table 2: Headword spelling comparison among dictionaries (adapted from Ellis 1993).

Table 2 presents a comparison between the different headwords selected by different dictionaries. The table summarises the two main tendencies described by Ellis, that is, a preference for Early West Saxon or Late West Saxon. In this respect, the DOE, as stated above, represents the most systematic approach towards the use of the latest available form, while Sweet makes use of the oldest spelling, even if the precise form is unattested. Unattested forms are presented in the chart by means of (*). Clark Hall and Boworth and Toller represent compromising solutions, though,

again, heading in different directions. Whereas Clark Hall aims at using the oldest spelling, he is more conservative than Sweet and adopts modern spellings where the potential oldest form is not attested in the texts. Bosworth and Toller attempt to make use of the latest spelling, but they are more unsystematic than the DOE. Differences can also be observed between Bosworth-Toller's and Bosworth's dictionaries. As has already been mentioned, Bosworth did not pose any load of prescriptivism to his authors when compiling the *Compendious Anglo-Saxon Dictionary*, and this reflected in a series of inconsistencies in the spellings selected, where forms include spellings with <e>, <i>, <y>, which correspond to different periods of the West Saxon dialect.

Leaving aside the question of spelling, there are other features affecting the headwords where the dictionaries present a diversity of approaches. While most dictionaries, whit the exception of the thesauri and some etymological dictionaries, are organised alphabetically, Sweet, includes an innovative mixed system. Whenever a word becomes the base of compounds or derivations, the derived elements are listed immediately before it, thus breaking the alphabetical order, which is resumed once the derivative paradigm of the word in question is completed. Consider the case in (3) where the headwords *scīr* (f.) and *scīr* (adj.) are non-consecutive entries:

Scīr *f.* office, administration; district, shire, diocese, parish.

~biscop *m.* bishop of a diocese.

~lett *n.* piece *or* measure of land.

~(e)mann, scīrig- *m.* official, steward; procurator; native of a district.

~gemot *n.* shire-mote

~gerefa *m.* judicial president of a shire, sheriff.

~gesceatt *n.* property of a see

~geþegen *m.* thane of shire

(e)wita *m.* chief man of shire.

scīr transparent, clear (weather); bright, glittering, white, brilliant; pure (wine); clear (voice); splendid.

~baso bright purple.

~e *av*. Brightly; clearly (*of* voice).

~ecg bright-edged.

~ham in bright armour.

~mæled with bright ormanents (sword)

~wered bright (light)

scīran declare, tell, speak…

As seen in example (3) the expected alphetical order is broken to include the lexical family of the noun *scīr* prior to the adjective *scīr*, and the same holds true for the lexical family of the adjective, which is displayed before the verb *scīran*.

3.3.2. Scope

To compare and assess quantitatively the information compiled in the reviewed dictionaries, the number of entries presented in a given range of headwords has been considered. For so doing, I have taken into consideration the range of words starting by *fe-* to words starting by *feo-*. The selection of items to compare has been done bearing in mind the restricted charater of the DOE in its present state. The quantitative summary of the entries found between the chosen spellings is displayed in table 1:

	Clark Hall	Sweet	DOE	Bosworth-Toller
Number of entries	224	127	185	173

Table 3: Number of entries per dictionary.

As can be seen, Clark Hall stands out as the most complete dictionary of the four under comparison in quantitative terms. This simply means that Clark Hall presents a greater number of headwords. It may not be the case that it displays more information, as some of the headwords may just be inflectional forms or simply variant spellings which refer the reader back to another headword. To show an

example of the numeric information presented in (2), I list some of the headwords contained in the selected range. Consider Figure 2:

Clark Hall	Sweet	DOE	Bosworth-Toller
fearr	*fearr*	*fearr*	*fearr*
ø	ø	*fearre-mearg*	ø
ø	*fear-hryþer*	*fear-hryþer*	ø
ø	ø	*fearrlic*	ø
feas	ø	ø	ø
fēasceaft	ø	*fēasceaft*	*féa-sceaft*
fēasceaftig	ø	*fēasceaftig*	*féa-sceaftig*
fēasceaftnes	ø	*fēasceaftnes*	ø
feast	ø	ø	ø

Figure 2: Number of headwords per dictionary.

As shown in Figure 2, both Clark Hall DOE display a similar number of headwords, which clearly outnumber the ones proposed in the other dictionaries. However, there are significant differences among them two. On the one hand, the DOE introduces three complex words based on the noun *fearr* 'beast of burden', which are not present in the other sources, with the only exception of *fear-hryþer* 'bull' which can also be found in Sweet dictionary. On the other hand, Clark Hall counts the terms *feas* and *feast* as headwords, but they are actually variant spellings of *fæs* 'fringe, border' and *fæst* 'firm, secure' respectively, and the only information displayed in these headwords is the reference to the canonical terms to which they are related.

3.3.3. Alternative spelling and cross-references

Closely connected to the number of headwords and their organisation is the question of spelling. Old English was not stable at any linguistic level, as would be the case to any language from which a scope of 400 years is selected for analysis. One of the levels that are not stable is spelling, as shown by the Present-Day English word

Thames whose evolution, as attested in the *Oxford English Dictionary* is shown in (3):

(3)

Temes – Temese
Temze – Temeze (Tamise)
Temys – Temmes(se) – Themes – Themys – Themise – Thamyse – Thamise
Thames

Because of this variation, the spelling information available in an Old English dictionary is a key element. Diatopic and diachronic variation constitute one of the major problems as regards lexicographical analysis of the Old English period. When several spellings are available for a given form, decisions have to be made regarding which lexeme is going to be considered as canonical and which ones are to be treated as alternative forms of the word. The dictionaries under observation also show differences in this respect, as proved by Figure 3. The first element for each dictionary stands for the canonical form of the verb, whereas the rest of elements are considered alternative forms to the canonical one:

Clark Hall	Sweet	DOE	Bosworth-Toller
fahnian, fægnian	*fægenian, fahnian, fægnian*	*fahnian, fægnian*	*fægnian, fagnian, fægnigan, fægenian*

Figure 2: Canonical and alternative spellings.

As was the case with headword organisation, the treatment of variant spellings is inconsistent and leads to problems of circularity. Consider the examples in (4) taken form Clark Hall:

(4)

a. *ābugan* (=on-) ‘bow, incline, bend, submit’
   onbugan ‘to bend; bow, submit, yield to’
b. *onbyhtscealc* = *ambihtscealc*
   ambihtscealc ‘functionary, retainer’
c. *oncierran* (e, i, y)
   oncigan (ei = ie)

As (4) shows, Clark Hall is inconsistent in the treatment of spelling variation. (4a) shows that the author acknowledges the fact that *ā-* and *on-* are variant forms of the same prefix. However despite providing that information in the entry for *ābugan*, he creates another headword for *onbugan*. However, in this second headword, no information is given as regards alternative spellings of the prefix. In both cases, a translation of the terms is provided. It should be remarked that the translations are not, nonetheless, identical. In (4b), however, Clark Hall identifies two alternative spellings and refers the reader to the second term to check its meaning. As in the previous case, the reference to the variation is unidirectional. No reference to the form *on-* is made under the headword *ambihtscealc*. Finally, (4c) proves inconsistencies as regards the choice of a standard form the headwords. This example refers to two consecutive headwords of the dictionary. In the first case, the selected spelling for the headword is the diphthong *-ie-*, which possesses the alternative forms *-e-*, *-i-*, and *-y-*. The following word however is presented with the canonical form *-i-* for which an alternative -ie- form is attested.

3.3.4. Phonology

A further area of differentiation lays on the treatment of vocalic quantity. Old English had seven simple vowels and four diphthongs, with their corresponding long variants, as presented in (5):

(5)

Short vowels: /i/ - /e/ - /æ/ - /o/ - /u/ - /a/ - /y/
Long vowels: /ī/ - /ē/ - /ǣ/ - /ō/ - /ū/ - /ā/ - /ȳ/
Short diphthongs: /ei/ - /io/ - /ea/ - /eo/
Long diphthongs: /ēi/ - /īo/ - /ēa/ - /ēo/

This phonological distinction is not, however, signalled in the original texts. Rather they are a more modern lexicographical representation, and as such, it refers different treatment by different authors. Vowel quantity in Old English is important to identify lexical items, as different vocalic length implies dealing with different words as shown in (6):

(6)

bær 'bare, naked, unclothed vs. *bǣr* 'a bier, handbarrow, litter'

Clark Hall and Sweet include information on vocalic quantity, and indicate vocalic length by means of a macron (¯) placed upon the affected vowel or upon the first element of a diphthong. The DOE also accounts for vocalic quantity in their headwors while, following the original texts, it does not account for it in the textual material included under the headword to assess the different forms of each word. Bosworth-Toller on their part, make use of the diachritic (´) with an ambiguous meaning. It sometimes denotes vocalic length while in some other occasions shows stress position, especially when distinguishing derived from compound lexemes. Take (7) as illustration:

(7)

bær 'bare, naked, open' *vs*. *bǽr* 'a bier, *feretrum*'
fór-tácen 'a *fore-token*' *vs*. *for-téah* 'misled, seduced'

This unsystematic, dual use of the symbol (´) leads to problems of interpretation of the word under observation. On the one hand, this dictionary does not always include this information in its entries and on the other, the information included may lead to misunderstandings regarding the position of the stress in the word, as we can see in example (8) where a comparison among the four dictionaries is carried out:

(8)

BT: *candel-leóht*.
SW: *candel-leoht*.
CH: *candel-lēoht*
DOE: *candel-lēoht*

What we see in example (8) is that, if the diachritic shows word stress, it is wrongly placed, as compounds in Old English are always stressed on the first element. If we consider it as a vowel quantity marker, there is a conflict with Sweet's (1898) proposal, while showing agreement with the other two compared works. This intuition is also confirmed by the treatement given to this word in other sources.

A difference in the way dictionaries are consulted emerges also in this respect. While traditional printed works must be carefully checked, electronic dictionaries can be consulted in a more careless manner. In electronic format, searches can be launched disregarding vowel length and making use of the short vowel. The results thrown by the DOE include both forms with long and short vowels.

3.3.5. Syntax

The information presented in the different entries regarding syntactic matters receives different treatments in the different dictionaries under revision. To begin with, Sweet (1976) is quite unsystematic in the display of syntactic information. As a general rule, not much information is provided as regards the transitive or

intransitive nature of a verb, especially in the case of weak verbs. The intransitive character of verbs is remarked in two main general contexts, namely, when the verb is strong or when there are semantic differences derived from the transitive or intransitive character. Consider the examples in (9):

(9)

Būgan *7 intr.* Bow dwon, stop; bend, swerve (sideways *or* up and down)…
Hwierfan *tr.* Turn,;change, transform; convert (to religion) || *intr.* Turn, revolve; return; wander, go , move; change.

In the case of weak verbs, syntactic information may or may not be present. Compare the cases in (10)

(10)

hwearfian *intr.* Turn, revolve, roll; wave (of banner); wander; change
hwierftlian revolve; wander.

Even if the verbs are closely connected, as indicated by the similarities in shape and meaning, the treatment they receive is different.

In the same vein, Bosworth-Toller are quite unsystematic as well. Verbal headwords in this dictionary may or may not provide the reader with information on the syntactic nature of the verb, whether transitive or intransitive. Alternatively, some headwords containing several separate definitions offer information regarding the complementation pattern in a specific use of the word. Compare the cases in (11):

(11)

a. **ge-wyrdan, werdan**; *p.* de; *pp.* ed; *v.trans. to hurt, injure, lædere, nocere*

b. **ge-wurþian;** *p.* ode, ade; *pp.* od, ad. *To distinguish, honour, adorn, celebrate, praise.*

c. **ge-wurþan;** he -wurþ; *subj. pres.* -wurþe, *pl.* -wurþon. **I.** *to be, become* […] **II.** *v. impers. cum acc. to happen, come to pass, come toghether, agree.*

Ge-wyrcan, wyrcean; *p.* -worhte, ðú -worhtest: *pp.* -worht. **I.** *To work, make, build, form, dispose, do, perform, celebrate, commit.* […] **II.** *to get by working, gain, obtain, merit.* […] **III.** *with gen.* [cf. wyrcan *with gen.*] For hwam nele mon him georne gewyrcan dryhtscipes *why will not man earnestly gain himself worship.*

On its part, the DOE does not provide the reader with explicit information on the transitive or intransitive character of verbs. Rather, in each of the meaning definitions, the textual material and exemplification is included for each headword, it refers the complementation pattern of some of the meanings is provided, referring the reader to the texts so that the syntactic structure of verbs can be inferred. Consider the case of *ēacnian*, presented in figure (3):

ēacnian

Vb., wk. 2

1. to increase

1.a. without expressed object: to increase, become greater in intensity

Rim 30: burgsele beofode, beorht hlifade, ellen **eacnade**, ead beacnade.

1.b. to increase, add to, augment (someone *acc.*)

Gen (L) 17.20: ofer Ismahel eac swilce ic gehyrde þe. efne ic hine bletsie & **geacnige** & swiþe ic hine gemanyfilde (Gen *geeacnige;* cf. Gn: *ecce benedicam ei et augebo et multiplicabo eum valde*).

2. to add (something)

ÆGram 107.3: þry eacan synd *met, pte, ce,* þe man **eacnað** on ledenspræce to sumum casum þises partes for gesceade oððe fægernysse.

3. to conceive

3.a. without expressed object

ÆLet 4 (SigeweardB) 576: *ecce uirgo concipiet & pariet filium & uocabitur nomen eius Emanuel* efne mægden sceal **eacnian** & acennen sunu & his nome bið icwædon God sylf is mid us.

3.b. figurative: to conceive anger (in one's heart)

Figure 3: Partial entry in the DOE, reflecting the complementation patterns of several meanings.

Finally, Clark Hall does not usually provide syntactic information, but he makes an exception on the matter when the complementation pattern of the verb, which sometimes depends on morphological matters, implies semantic changes, as can be seen in (12):

(12)

(ge)openian: 'to open, open up, disclose, declare, reveal, expound'. (*ge*) intr.: 'become manifest'. (ge) be open to

In this case, the use of the prefix *ge-* alters the semantics of the verb and consequently modifies its complementation pattern.

3.3.6. Morphology

The subject of morphology, both derivative and inflectional is treated differently in the various dictionaries reviewed. As regards inflectional morphology, the information presented is not standardised, not only in the comparative study of the dictionaries but in each dictionary as well. The current subsection focuses on the treatment of the major lexical classes.

Let us consider first the case of Clark Hall's (1996) dictionary. For the nominal paradigm, he provides the gender of the word, whether masculine (m), neuter (n) or feminine (f). In some words, as *bīwist* (mf) 'sustenance, food', he shows a double gender and even a triple gender as in *amber* (mfn) 'vessel, pail, cask'. However, there is no indication on whether the nouns follow the weak or the strong declension or any of the other minor declensions as the *i*-stem or the *u*-stem declension. Only in a very reduced number of occasions Clark Hall does add and

indication of some inflectional form which could be informative of the paradigm a word should be adscribed to. This is shown by *bæl-dracu* 'violence of fire' which appears in the dictionary as shown in (13):

(13)

bæl-dracu f (ds. dræce) *violence of fire*. PH270

In the case of verbs, a distinction has to be made as regards weak and strong verbs. Weak verbs do not show any indication of subclass distribution, while their consideration as weak verbs is seldom made explicit. Compare the cases in (14):

(14)

traht-ian: *To treat, commenton, expound, consider*.
treowan: trans. w.d. *to believe, trust in*

Regarding strong verbs, information on the lexical subclass, although present, is neither exhaustive nor systematic. Consider the cases in (15):

(15)

a. **lūtan[2]:** *to bend, stop, decline* ; **ðringan[3]:** *to press, squeeze, crowd upon*.
b. **berīdan[1]:** *to ride round, surround, besiege*.
c. **forbeodan:** *to forbid, prohibit*

Example (15) shows that the lexical class assigment is present in underived strong verbs, whereas it is unsystematic in the case of derived verbs. This manner of proceeding complicates the retrieval of information as it is often a must to go back to the underived verb to check its inflectional morphology.

A second issue with the presentation of the inflectional forms of verbs, is that they are not usually presented as part of the lexical information of the infinitives, although some exceptions may be found, as in the entry for *unnan* ‘to grant’ shown in (16):

(16)

unnan: pres. 1 sing. an(n), on(n), pl. unnon; pret. sg. ūðe, pp. (ge)unnen. swv. w. d. pers and g. thing. *to grant, allow*

In the case of (16) a lot of inflectional information is provided. Some of the finite forms of the verb are acknowledged, as well as the fact that this particular strong verb presents some weak forms in the preterite tense. Moreover, syntactic information is provided, stating the complementation pattern of the verb. It is followed by a dative if the complement is a person, but followed by genitive if the object is a thing.

The usual treatment of inflectional form, however is to treat past participles and some finite forms of the verbs as proper headwords, although the only information they display is the inflectional form they represent and the infitive form of the verb to which they belong. Take (17) as illustrative:

(17)

***nile=nyle*:** pres. 1 sg. of **nyllan*.
wǣron: pret. pl. of wesan.
coren: pp of cēosan.

Finally, adverbs are sometimes marked as such by mean of the contraction *adv*., as in *deope* (adv.) ‘deeply, thoroughly’, but in most occasions, no reference to the lexical class is provided. In the same vein, adjectives are not usually marked, and they have to be interpreted through the translation provided. The only cases where

adjectival forms are signalled occur when they are homographs with a word from a different lexical class. A sample is shown in (18):

(18)

unrihthæmed: I. n. *fornication, adultery*. **II.** adj. *adulterous*
unseht: I. mfn. *Discord, disagreement, quarrel*. **II.** adj. *not agreed, hostile*.

Sweet's (1896) account of morphological information is also divergent and requires some comment. As for nouns, he only states the gender of the word, whether masculine, neuter or feminine, and only sometimes he adds inflectional forms, with or without indication of case and number. Examples are provided in (19):

(19)

lefung *f*. paralysis; **nīed-clamm** *m*. distress, need
sweostor: sister. pl. sweostra, -u

In the case of verbs, contrary to Clark Hall's (1996) general practice, weak verbs are indicated by means of the contraction *wd* (which stands for *weak declension*), although they are not assigned subclass. Strong verbs, however, are assigned subclass (but see chapter 2 on Sweet´s classification). Inflectional forms are not usually shown, unless they show some degree of irregularity. Compare the cases in (20):

(20)

a. **ridan** 6 *intr*. ride.
b. **ceosan**, *pret*. *pl*. curon, *ptc*. coren 7 choose.

In the examples above, (20a) shows a prototypical entry of a strong verb, whereas (20b) displays a verb whose citation forms are irregular. In this particular case, the

preterite plural and the past participle are affected by rhotacism. Mind that the preterite present form, being regular, is not displayed in the entry of *cēosan*.

Considering adjectives and adverbs, Sweet makes use of the same standard as Clark Hall. Adjectives are not marked unless coincidental in form with another lexical class, whereas adverbs are always marked by means of *adv*. See example (21):

(21)

hrēmig. Exulting
hrēow. I. *f.* penitence, repentance. **II. hrēo(w).** *aj*. sad, repentant
grame. adv. *angrily, fiercely, cruelly.*

Bosworth-Toller on their part are more systematic in their presentation of inflectional morphology. Apart from indicating the grammatical gender of nouns, they add one or more inflectional endings to make clear the declension a particular word should be assigned. Compare the words *fengel* and *feoh* shown in (22).

(22)

a. **fengel**, es; m. A prince; *princeps*.
b. **FEOH,** fioh; *gen*. feós; *dat*. feó; n. **I**. Cattle, living animals. **II**. Cattle being used in early times as a medium of exchange, hence *Money, value, price*. **III**. As property chiefly consisted of cattle, hence *Goods, property, riches, wealth*. **IV**. The Anglo-Saxon Rune ᚠ.

As for verbs, they all include the standard citation forms; preterite and past participle for weak verbs and preterite singular, preterite plural and past participle in the case of strong verbs. In the case of strong verbs, they usually display some attested inflectional spelling and, afterwards, the citation forms. This is shown in (23) where (23a) stands for a weak verb and (23b) for a strong one:

(23)

a. **dǽlan**; *p*. de; *pp*. ed. To divide, separate, distribute, bestow, spend, dispense.
b. **BEÓDAN**, biódan;, ic beóde, bióde, þu beódest, býtst, býst, he beódeþ, být, *pl*. beódaþ; *p*. He beád, þu bude, *pl*. budon; *pp*. boden; *v*. *trans*. **I**. *To command, bid, order*; **II**. *To announce, proclaim, inspire*; **III**. *To offer, give, grant*.

The other major lexical classes, that is adjectives and adverbs, are always marked by means of the contractions *adj*. and *adv*. as shown by (24):

(24)

a. **dígol**, dýgol, diógol; *gen*. *m*. *n*. digles, *f*. digolre; *def*., *nom*. *m*. dígla; *f*. *m*. digle; *adj*. *Secret, hidden, private, dark, obscure.*
b. **ge-hú**; *adv*. *In any manner*.

The only exceptions to this general rule happen when the dictionary entry corresponds to alternative spellings, in which case the lexical class of the word is not provided. Rather a definition of the word is given along with a reference to the canonical form where the lexical class is made explicit. Take (25) as illustration:

(25)

disg: *folish* v. Dysig.
disig: *folly* v. Dysig.
dysig *adj*. DIZZY, *folish, unwise, stupid*.

Finally, the DOE offers a more coherent account of the morphology of each word. For nouns, they specify gender and class (1 for strong declensions and 2 for weak ones). This is shown in Figure 4 below:

a. dǣd-weorc

Noun n., cl. 1

Att. sp.: dædweorce

1 occ. (in poetry)

deed, action; literally ‘deed-work’

Ex 575: hofon hereþreatas hlude stefne, for þam dædweorce drihten heredon, weras wuldres sang (emendation to dægweorce has been suggested).

b. bēhfþ

Noun, f., cl. 2

Att. sp.: with late wk. inflect.: behefðen (xii)

2 occ.

want, need; lichamlice behefþa ‘bodily needs’

LS 22 (InFestisSMarie) 5: ac Martha beswanc & bestuddede þa lichamlice behefðen (cf. Lc 10:40 Martha autem satagebat circa frequens ministerium).

LS 22 (InFestisSMarie) 80: Martha swanc & becarcade to geforðigene þan Hælende & his þeowen þa lichamlice behefðen (cf. RADULF.ESC. Hom.assumpt. 647A: ista laborat ad exhibenda indigenti omnia humanitatis officia).

Figure 4: Example of an entry giving full morphological information, in the DOE

After that basic information, they add a list of attested spellings, although they show no indication of the inflectional form they belong to. Following with the examples above, the attested spellings for the words in (25) are *dædweorce* and *behðe* respectively.

As for verbs, they indicate their class, either weak or strong, and sub-class, along with the representative citation forms. Furthermore, as it is the standard procedure of the dictionary they provide all the attested spellings. Consider the cases Figure 5 where the entries of a weak and a strong verb are displayed:

a. bǣtan

Vb., wk. 1

Att. sp.: bætan || bætte || bættan.

4 occ.

It is not certain that the uses with widely differing senses are all of the same verb.

1.a. to bridle; here in the sense ‘to bridle and saddle’ rendering sternere ‘to provide (a horse) with a cloth for the rider’

GenA 2867: ongan þa his esolas bætan gamolferhð goldes brytta, heht hine geonge twegen men mid siðian (cf. Gn 22:3 Abraham ... stravit asinum suum).

1.b. to spread (something) glossing sternere ‘to spread out’

ClGl 3 204: strauerunt bættan (perh. from Mt 21:8 plurima autem turba straverunt vestimenta sua in via [or take as corrupt form of brǣdan2]).

2. a sailing term of uncertain meaning; Modern English uses (see OED2 beat v.1 sense 19) suggest a sense ‘to strive against a contrary wind’; the Modern English verb is intransitive in this sense; the OE occurrence has the ‘contrary wind’ in the genitive; ‘to make fast’ (of a ship) has also been suggested

Bo 41.144.31: gif he ær þweores windes bætte, warenað he hine wið ðæt weder.

3. to bait, to set with animals upon (someone acc. or dat.)

ÆLS (Agatha) 84: ða andwyrde Agathes unforht þam deman, gif ðu mid wilddeorum me nu bætan wylt, hi beoð sona handtame, þurh þæs hælendes naman.

Lat. equiv. in MS: sternere

See also: gebǣtan, ymbbǣtan; gebǣte, gebǣtel; bǣting; cf. bēatan, bītan

MED baiten. OED2 bait v.1 Cf. OED2 beat v.1

b. biddan

Vb., st. 5

Att. sp.: biddan, biddon, biddæn, bidden; bidda (Li, Ru) | byddan.

bidde, bide (pres.ind. 1st sg., LS 20), biddæ (xii), biddu (PsGlA, Ch 1283), biddo, bido (DurRit, Ru); bid (pres.ind. 1st sg.) | bydde.

bidst, bitst, bitsð (Lit 4.9), bist (xi-xii); biddest (GD MS C), bidest, bidæst (ÆHom MS B) | bydst (WSGosp MS H), bytst, byst (Solil); byddest.

bitt, bit, bidt; biddeþ, biddeð, biddeth, bideþ (GD, HomU 19), bideð; biddað (pres.ind. 3rd sg., GD MS C); biddes (Li); bid (NicD) | bytt, byt.

biddaþ, biddað, bidaþ (Ru), bidað (HomM 5, WSGosp MS Cp), biddad (HyGl 3 MS V); biddas (DurRit, Li, Ru), bidas (Li); biddæþ (xii), biddæð (xii); biddeð; biddes (Li); bidde (pres.ind.pl.) | byddað; byddeð (WSGosp MS H), bydðed (WSGosp MS H).

bidde | bydde (ChrodR 1).

bidden, biddan (pres.subj.pl.), biddon; bidde (pres.subj.pl.).

bide, bidde (imp.sg.), biddu (MCharm 11), bid (BenRGl).

biddaþ, biddað, bidað (WSGosp MS Cp); biddas (Li); biddæð (xii); biddeð (imp.pl.) | byddað; byddeð (imp.pl., WSGosp MS H).

bæd, bædd | baed (Li), baedd (Li) | bæt (AldV 1, ÆHom 20 MS C), bæð (WSGosp MS H) | będ | bed, bedd (Ru) | bead (xii) | beædd (Li) | beed (MS beed hine den over eras. of pret.ind.pl. form, WSGosp MS H).

bæde | bede; bedu (LS 5) | beede (WSGosp MS R).

bædon, bædan, bædun, bæden (pret.ind.pl., xii); bædo (ChronE); bæde (pret.ind.pl.) | będon (ChronA) | bedon, beðon (Li), bedun, bedæn (xii), beden (pret.ind.pl., xii) | beadon (ChronE), beaden (ChronE).

bæde | będe (Bede MS O) | bede.

bæden, bædon (pret.subj.pl.) | beden, bedon (pret.subj.pl., xii).

to biddanne, to biddenne, to biddene; to biddende (infl.inf., ÆCHom MS A).

biddende, bidende, biddenda; biddend (pres.part., m.nom.sg., Li); biddynde (PsGlC) | byddende || biddendne || biddendes || biddendum ||

biddende, biddenda, biddande (HyGl 3) || biddendra || biddendum || biddendan (wk., acc./dat. sg.).

beden || bedene || bedenan.

Late: beddan (BenRW) || byddaþ (PeriD) || bede (pres.subj.sg., ÆGram MS W).

Abbrev. forms in gloss: bidd' (DurRit), bid' (DurRit).

ca. 3200 occ.

Figure 5: full entry in the DOE, including all the attested spellings and their location in the corpus.

To finish off, adjectives and adverbs are always marked, and their attested spellings shown, as can be seen in (26):

(26)

a. bæsten

   Adj.

   Att. sp.: bæstenum.

b. dēofollīce

   Adv.

   Att. sp.: deofollice, deofolice; deoflice.

When considering derivational morphology, the different authors have also opted for a series of variant methodological decisions. Clark Hall follows an alphabetical order and includes the derived predicates in ther corresponding slot, without further information concerning the base of derivation. Sweet, however, despite following alphabetical orderding, accounts for the base of derivation of the complex word. Bosworth-Toller also follow alphabetical order, but in their case, the prefixes that can be attached to a given base are provided in the entry for the underived term as well. The DOE follows this same metholodogy. Unfortunately, some of the derived lexemes, although anticipated in the underived predicate, cannot yet be retrieved because of the unfinished state of the dictionary. Example (27) summarises the way derivational morphology is treated in these works:

(27)

Clark Hall: **ymbfaran**[6] *to surround*, AO80: *travel round*, GD490.
Sweet: **ymb-snidenness** *f.* circumcision. [ymbsnīþan].
Bosworth and Toller: **be-swícan,** *to deceive, entice, seduce*.
swícan *to move about, wander*; *to move away, depart, escape*; *to desist from, cease from*; *to deceive*; *to fail in one's duty*. DER. *á-*, *be-*, *ge-*.

diPaolo Healy *et al.*: **be-faran**, Vb., st. 6, to go, proceed; with reflexive: to betake oneself, go.
faran, vb., st. 6, expressing movement: to go, travel, proceed, make one's way.
[…]
See also ge-, ā-, be-, for-, forþ-, full-, geond-, of-, ofer-, oþ-, þurh-, tō-, wiþ-, ymb-faran.

Example (27) shows the different ways in which derived elements are treated, with or without crossed references to other entries in the dictionary; in the case of The DOE all potential derivational forms are attested as in *of-faran*, *ofer-faran* or *oþ-faran*, although the entries for those lexemes have not yet been developed.

3.3.7. Meaning definition

Perhaps the point where the methological criteria behind the compilation of each of the dictionaries under comparison is more similar, is semantics. Basically, all the dictionaries have opted to follow an order based on prototypicality and frequency. Consider the cases in (28):

(28)

Clark Hall: **+bed I.** n. *prayer, supplication, Æ,* CP *religious ordinance, service, Æ*.

Sweet: ge**bed** *n*. prayer; religious ordinance *or* ceremony.

Bosworth-Toller: **BÉD**, es; *nom. acc. pl.* bédu, bédo; *n. a prayer, supplication, religious worship*.

DOE: ge-bed, Noun, n., cl. 1. 1. praying, prayer (normally addressed to God), 1.a. prayer, supplication; frequently in association with other good deeds.

As we can see in (28), both Sweet and The DOE present the headword *béd* with the pre-verbal element *ge-*, whereas Clark Hall and Bosworth-Toller do not. However, Clark Hall presents the headword with the symbol (+), which suggests that the headword may appear in more instances with some other pre-verbal elements. On the part of semantics, it can be seen that the DOE gives the most exhaustive definition of the four dictionaries, giving a sense that goes further from the strictly religious meaning the headword may have in the other dictionaries.

The organisation of the meanings also offers matters for discussion. Consider the cases in (29):

(29)

Clark Hall> **behēawan**. to cut, chip, chop, beat; cut off from, deprive of

Bosworth-Toller: **behēawan.** to beat, bruise, hew or cut off, to separate from, deprive of.

DOE: 1.a. figurative: to strike (as with blows)

2. to cut (someone *acc.*) off.

2.a. to deprive (someone *acc.*) of (something *dat. / instr.*); *ealdre beheawan* ‘to deprive of life, kill’

2.b. *heafde beheawan* ‘to deprive of one's head, behead’

Even if the meanings provided are similar, differences arise regarding the organisation, distribution and importance given to each meaning. Clark Hall makes

use of a semi-column to indicate a change in the semantic domain or use of the word. Bosworth-Toller, however, treats every meaning in the same manner, and provides no separation between the different uses of a word if the meanings are so similar so as not to postulate a sencond semantic entry. The DOE, on its part, displays a clear separation between each specific meaning, and even shows more concrete, detailed distinctions. In the example above the verb belong in the semantic domain of the *hitting* verbs. Two general ideas stand out, hitting with a blunt or a sharp object. Clark Hall lists first the meanings having to do with cutting, or hitting with a sharp object, while offers just *beat* as an option for hitting with a blunt instrument. A second set of meanings is provided, separated from the rest, and where the expreses idea is thant of *cutting to the point of separation.* All these meanings are also put forward by Bosworth-Toller. However, they offer a list of synonims of hit at the beginning, to later introduce the concept of cutting and the more specific notion of cutting to the point of separation. Nonetheless, no distinctions are provided by means of typographic signs. The DOE, includes under the headword *behēawan* two different meanings; one for hitting or, to be precise, striking, and one for cutting. Besides, this separation, it states *figurative use* in the case of the meaning *strike*, a piece of information which is completely omitted in the other dictionaries. The second meaning, that of cutting, is further divided from a general "cut off" meaning, to a more specific, cut to the point of separation, to finish with the more specific meaning of beheading.

As can be seen, the compared dictionaries show differences in the presentation of the semantics of headwords, even if the displayed meanings are the same. In some occations, however, the meanings provided are not so coincidental as in (29). Consider the case in (30):

(30)

Clark Hall: **geondwadan**. to know thoroughly, be versed in.

Bosworth-Toller: **geondwadan.** go through a subject, make oneself acquainted with, study.

DOE: **geond-wadan.** to know thoroughly, be versed in (a subject *acc.*)

In (30) Clark Hall and the DOE coincide in supplying the same meaning to the word, providing the reader with a stative definition of the verb. However, Bosworth-Toller, defines the verb as an activity, the process needed to become versed in a subject. Even if the semantic domain is similar, the definitions presented offer conceptual distinctions divergent perspectives in approaching the semantics of this particular verb.

3.3.8. Etymology

As in all the other areas of comparison, divergences arise in the quantity and quality of the information on etymology provided in the dictionaries.

Sweet does not take into consideration the origins of the words, neither from a morphological nor from a semantic perspective. No reference to this aspect is made in any of his dictionary entries. Clark Hall does not consider special attention to this aspect either, but does acknowledge some etymological information in some entries, as it is the case with *ǣren*, shown in (31):

(31)

ǣren I. made of brass, brazen, *Æ*, AO, CP: tinkling? [ār; cp. *Ger*. Ehern] **II**. Oar-propelled, GD347.

Bosworth-Toller do include etymological information in a more systematical manner. Most of the words include information of older stages of the language as well as other cognate terms found in other Germanic languages, as shown by (32):

(32)

sāre *sorely, grievously, bitterly* [O. Frs. *sēre*: O. Sax. O. H. Ger. *sēro dolenter*: Ger. *sehr*.].

Cin, cyn, e; *f. The* chin; mentum […] [O. Sax *kinni, n*: O. Frs. *kin, ken*: Dut. *kin, f*: Ger. M. H. Ger. *kinn,* n: O. H. Ger. *kinni, n*: Goth. *kinnus, f. the cheek*: *Dan. kind, m. f.*: *Swed. kind, f*: *Icel. kinn, f*: Lat. *gena*: *Sask. hanu, m. f. the jaw*].

Finally, the DOE provides etymological information and cognate terms in other Indo-European languages as shown by (33):

(33)

ge-feallan 'without expressed object'. Lat. equiv. in MS: +cadere, corruere, decidere, defluere, descendere, excidere, incidere, irruere, labare, pluere, procidere, procumbere, proicere, prosterni, provolvere, ruere
MED *ifallen*. OED2 *ifalle, yfalle*.

Besides, these authors resort to the *Oxford English Dictionary* to asses the evolution of the word in later stages of English.

3.3.9. Textual evidence

The final aspect of comparison is the treatment of the textual evidence that supports the inclusion of a given heardword or form in the dictionary.

Again, Sweet does not provide the reader with that information. Clark Hall includes the acronym of the text that accounts for the existence of the word at the end of each entry. Consider (34):

(34)

niðerhrēosende (y) *falling down*, Æ.

Bosworth-Toller and the DOE are the two dictionaries that best exemplify lexical elements. Concerning Bosworth-Toller, they include short samples of the texts containing the word, along with the acronym of the particular text, and more specific information detailing the precise part where the word is to be found. Take (35) into consideration:

(35)

BROC, es; m.? A brock, badger; taxo = tassus [tasso *It*: taisson *Fr*.], meles:- Broc *taxo* vel *melus*, Wrt. Voc. 22, 53. Sum fyðerfēte nȳten is, ðæt we nemnaþ taxonem, ðæt ys broc on Englisc *there is a four-footed animal, which we name taxonem, that is brock in English*, Med/ ex Quadr. 1, 2: Lchdm. i. 326, 12 [*Wyc*. brok: *Laym*. brockes, *pl*: *Dan*. brok: *Icel*. brokkr, *m*: *Wel*. *Corn*. broch: *IR*. broc, *m*: *Gæl*. broc, bruic, *m*: *Manx* broc, *m*: Armor. broc'h, *m*]

The DOE follows a similar structure, but include a textual reference not only for the headword, but also for each of the attested spelling, thus providing the reader with a more clear and accessible picture of the state of affairs. Consider (36) as a sample:

(36)

earm-bēag
Noun (m., cl. 1)
Att. sp.: earmbeag, earmbeah | ermboeg || armbages (m.nom.pl., WerdGlA) || earmbeaga
6 occ. (in glosses and Beo)
arm-band, bracelet
Beo 2756: geseah ða sigehreðig ... maððumsigla fealo, gold glitinian grunde getenge, wundur on wealle ... þær wæs helm monig eald ond omig, earmbeaga fela searwum gesæled.
HlGl D410: dextrocerium .i. brachiale, armillum **earmbeag**.
AntGl 6 791: dextrochirium brad **earmbeah**.

CollGl 11 24: dextrocerium **earmbeag**.

LdGl 19.43: armilla **ermboeg**.

WerdGlA 4.29: dextralia **armbages**.

As way of summarising, along this section, I have carried out a comparison of different lexicographical sources on Old English. The comparative analysis has outlined the strengths and shortages of each proposal, and has remarked the importance of the methodological principles underlying each lexicographer's decision. Given the importance of the methodological statements when dealing with any kind of research but, more particularly, when conducting a lexicographical investigation, the following section concentrates on describing the theoretical and practical aspects that govern this thesis. More specifically, these coming sections describe the way the data are going to be treated and the preliminary work that has to be done when the procedure to be applied makes use of the new techonlogies and databases rather than being based on the manual work that lies behind traditional lexicography.

3.4. Relevance, goals and scope of the research

Finally, as the thesis here undertaken is based upon corpus research, no indication of vowel length will be included in the final lemmatas, as that information is not included in the textual material used.

This research deals with the strong verbs of Old English. Its aim is to provide an inventory of lemmas of strong verbs that is based on the textual evidence provided by The Dictionary of Old English Corpus.

The lemmatisation of strong verbs in the Old English language, with textual evidence. There are several dictionaries focused on Old English An Anglo-Saxon Dictionary, by Bosworth and Toller (1973), A Concise Anglo-Saxon Dictionary, by Clark Hall (1996) and the Dictionary of Old English by diPaolo Healey et al. (2003) yet, as shown in chapter 2 they show several inconsistencies, not to say discrepancies, as regards not only their general organisation but, more also and most

importantly, in what respects to the selection of their headwords or lemmas. Thus, the main aim of this research is to provide an alternative system of lemmatisation of the Old English strong verbs that may be applied to the online database The Grid. This thesis will, as well, provide an automatised way to analyse and label inflectional forms in an easier and standardised manner. The goal is to open the way for a comprehensive analysis of the unlemmatised corpus, and not only for a study of verb forms.

In pursuing and reaching this final aim, several minor objectives need to be reached. First, turning the corpus into an indexed and searchable tool. The corpus used in the database to analyse the data is taken from the Dictionary of Old English Corpus (henceforward DOEC). The Dictionary of Old English (henceforward DOE) gives a thorough definition of the vocabulary of the first six centuries of the English language, through computer-based tools. The DOE is a complementation of the Middle English Dictionary, covering the period between the years 1100-1500, and the Oxford English Dictionary. The combination of the three dictionaries provides a full description of the vocabulary of English.

So far, several research tools have been published: the Dictionary of Old English Corpus on the World Wide Web, the DOE: A to G online, the DOE: A to G on CD-ROM, the fascicles for the letters A-G on microfiche, and an online bibliography of Old English texts and Latin sources cited in the DOE. In addition to this, eight of the twenty-two letters of the Old English alphabet have been published.

The DOE in itself is based on a computerised Corpus including one copy of each surviving text in Old English. The body of the corpus includes a compilation of records on parchment, carved in stone and inscribed in jewellery. The texts are classified in different categories, including prose, poetry, glosses to Latin texts and inscriptions. Talking about prose, it can be found texts devoted to saints' lives, sermons, biblical translations, penitential writings, laws, charters and wills, records, chronicles, a set of tables for computing the moveable feasts of the Church calendar and for astrological calculations, medical texts, prognostics, charms, and cryptograms.

From the beginning, this project has used innovative methods and technologies. One of the characteristics that take apart the DOE from the rest of

dictionaries is that the corpus is the one that determined the headwords, whilst the rest of dictionaries are strongly based on the previous lexicographical works. The reliance on the corpus is done thanks to a comprehensive analysis of the records of Old English, which marks the difference between the DOE and earlier works. This project has been influenced by the new standards established in modern lexicography, through the provision of lists of all attested spellings, grammatical information, and meaning divisions for each entry.

This corpus has been indexed to fit the purpose of this work. For so doing, a concordance of all the words contained in the corpus has been carried out, which includes fifty characters of the surrounding text on either side of the concorded word. A sample of the concordance is offered in Figure 1:

Prefield	**Concorded Term**	**PostField**
	Eala	ðu cleric ne wana ðu æfre wexbreda fram sidan. fo
Eala	ðu	cleric ne wana ðu æfre wexbreda fram sidan. forfl
Eala ðu	cleric	ne wana ðu æfre wexbreda fram sidan. forfleoh wes
Eala ðu cleric	ne	wana ðu æfre wexbreda fram sidan. forfleoh wesan
Eala ðu cleric ne	wana	ðu æfre wexbreda fram sidan. forfleoh wesan ealdo
Eala ðu cleric ne wana	ðu	æfre wexbreda fram sidan. forfleoh wesan ealdor s
Eala ðu cleric ne wana ðu	æfre	wexbreda fram sidan. forfleoh wesan ealdor sliden
Eala ðu cleric ne wana ðu æfre	wexbreda	fram sidan. forfleoh wesan ealdor slidendes plega
Eala ðu cleric ne wana ðu æfre wexbreda	fram	sidan. forfleoh wesan ealdor slidendes plegan

	þæt
Eala ðu cleric ne wana ðu sidan æfre wexbreda fram	. forfleoh wesan ealdor slidendes plegan þæt ne bl
la ðu cleric ne wana ðu æfre wexbreda fram sidan.	forfleoh wesan ealdor slidendes plegan þæt ne blissige fræ
ric ne wana ðu æfre wexbreda fram sidan. forfleoh	wesan ealdor slidendes plegan þæt ne blissige fræte bær
wana ðu æfre wexbreda fram sidan. forfleoh wesan	ealdor slidendes plegan þæt ne blissige fræte bær ne hel
u æfre wexbreda fram sidan. forfleoh wesan ealdor	slidendes plegan þæt ne blissige fræte bær ne helle ealdor.
breda fram sidan. forfleoh wesan ealdor slidendes	plegan þæt ne blissige fræte bær ne helle ealdor. gewin
ram sidan. forfleoh wesan ealdor slidendes plegan	þæt ne blissige fræte bær ne helle ealdor. gewin þæt
sidan. forfleoh wesan ealdor slidendes plegan þæt	ne blissige fræte bær ne helle ealdor. gewin þæt sy
an. forfleoh wesan ealdor slidendes plegan þæt ne	blissige fræte bær ne helle ealdor. gewin þæt sy halig eal
eoh wesan ealdor slidendes plegan þæt ne blissige	fræte bær ne helle ealdor. gewin þæt sy halig ealdor ne

Figure 6: Example of concordance in the corpus.

This concordance isolates every word in the corpus, and allows for its conversion into database format. The concorded terms are considered to constitute the list of inflectional forms which are to be used as the source for the organisation of the strong verb lemmas. This identification of individual forms allows for their indexation and future observation and analysis. The indexation of the corpus is the necessary stage for reaching the second goal, that is, carrying out a comprehensive analysis of the unlemmatized words A-Y and not a partial one as it is the case with the DOE. Starting from the identification of a set of underived strong verbs, the correct codification of searches allows for the identification of complex, derived

versions of a particular verb without having to follow an alphabetical order, which would imply a circular way of working. This way of facing the corpus minimises the amount of analysis that needs to be carried out. In fact, only the inflectional form of simple verbs need to be analysed. Results on the derived forms will be thrown by the database, provided the correct search coding. Take example (37) as illustration:

(37)

BASIC FORM: *standan*.
DERIVED FORMS: *ætstandan*, *ānstandan*, *āstandan*, *bestandan*, *forestandan*, *forstandan*, *framstandan*, *instandan*, *ofergestandan*, *oferstandan*, *ofstandan*, *ongeānstandan*, *onstandan*, *oþstandan*, *tōstandan*, *þurhstandan*, *undergestandan*, *understandan*, *wiþerstandan*, *wiþstandan*, *ymbstandan*.

The indexation of the corpus will permit the identification of the inflectional forms of the verb *standan* 'to stand'. From that moment onwards, search codes may be added to look for preverbal elements, allowing for the identification of the inflectional forms of the derived verbs above presented while enabling the finding of possible alternative spellings of the bound morphemes. This is in fact the third goal of this research, to contribute to the study of morphophonological alternations in the Old English lexicon. Through this initial approach to the lemmatisation of the corpus, patterns of spelling will be identified that will contribute to a better, more accurate system of searches when facing the lemmatisation of other categories. Search procedure, codes and findings will be dealt with in section 3.5 below.

As stated at the beginning of theis section, this research is limited to the study of the seven classes of Old English strong verbs. That means that weak verbs, as well as anomalous verbs, fall out of reach in this work. Furthermore, the non-canonical strong verbs, the so-called contracted verbs have also been disregarded. The reason for their exclusion is that their inflectional paradigms are not fully standardised and, for a preliminary approach like the one here intended, that fact would complicate matters much more than desired.

5. SUMMARY, CONCLUSIONS AND FURTHER RESEARCH.

5.1 Introduction

This chapter puts an end to this doctoral dissertation by offering a summary of the research process and its findings. It is divided into three main sections. Section 5.2 offers a summary and synthesises the analytical steps that have been taken, with special emphasis on the contribution of this work to the studies on Old English lexicography. Section 5.3 discusses the main conclusions on both the qualitative and the quantitative side. To round off, some final remarks are made in Section 5.4, where the lines of future research indicated by this undertaking are also discussed.

5.2 Summary

This dissertation has concentrated on the identification of the inflectional forms of strong verbs of Old English in a lexical database with the double aim of developing an electronic search tool and establishing a set of strong verbs lemmas.

Chapter 2 constitutes a review of lexicology, including the form that words adopt, their function and meaning, along with the processes through which new words are coined. A general review of the notion of lexicalisation and of word formation processes, from the global processes of compounding, derivation and inflection, to the more specific concepts of coinage or metalinguistic citation has been carried out. As the focus of study is Old English, the second part of the chapter constitutes a description of its morphophonemic features together with the status of its verb system. Given that the main goal of this dissertation is the establishment of a set of strong verb lemmas, a historical review of the evolution of the strong verb system from Indo-European to Old English was in point in this chapter. Thus the meaning, changes, and impact of *ablaut* in the different historical stages have been described. Finally, the Old English paradigms have been considered, including the description of the weak and strong verb forms as well as the role that strong verbs

played in the derivation of Old English. In describing the strong verb system, the different classifications of verb classes put forward by Clark Hall (1896), Krygier (1994), Levin (1964) and Sweet (1896) have been reviewed. Also, the description of the verb system involved the discussion of inflectional endings and their expected phonemic evolutions. To conclude this part of the study, the derivation of words based on the strong verbs was also addressed, including explicit derivation (prefixation, suffixation and compounding) and opaque derivation, as is the case with zero derivation.

Chapter 3 engages in the description of lexicography from a double perspective. First, a historical evolution from the traditional lexicographical methods, to the current situation, where technology and the digital evolution have played and are playing a remarkable role in the situation of the discipline. But, considering that traditional lexicography has still a great importance, especially in a language like Old English, this chapter also devotes special attention to the most important traditional lexicographical works focused on this stage of the English language. Lexicography, as an empirical science, focuses on the creation of dictionaries. A description of the different types of dictionaries, their motivations and the manner in which they treat and organise the lexical material has been provided. Secondly, the advances in the discipline, the incorporation ot technological devices and tools and the change from dictionaries to dababases has also been put forward. But although the evolution is constant and the impact of digital technologies is growing considerably, the works and procedures of traditional lexicographers cannot be dismissed. In acknowledging the relevance of traditional lexicography for Old English, a comparative study of four dictionaries has been carried out, which has allowed me to put forward its strengths and weaknesses, by paying attention to the different approaches with which they were created and the different methodological decisions that were taken during their creation. The dictionaries selected for this comparison were *An Anglo-Saxon Dictionary* (Bosworth and Toller 1973), *The student's Dictionary of Anglo-Saxon* (Sweet 1976), *A Concise Anglo-Saxon Dictionary* (Clark Hall 1996), and *The Dictionary of Old English in Electronic Form A-G* (Healey *et al*. 2008). The comparison included areas such as headword spelling, phonology, syntax, semantics and etymology among other. Having carried out the

contrastive analysis of the dictionaries, the final part of the chapter was devoted to the explanation of the scope, goals and contribution of the current research to the study of Old English lexicography. The last part of the chapter is devoted to the description of the methodological procedures underlining this work. The analytical steps go from the establishment of a digital corpus, its indexation and the development of searching to the processing of the retrieved results. Also, Chapter 3 puts forward some of the most relevant problems found during the research as well as the solutions put forward to overcome them including the establishment of restricting filtering codes and the reorganisation of prefixes into a system where circularity has been dismissed.

Chapter 4 describes the results of the analysis, which are displayed in four different domains. First, the list of inflectional forms identified for each of the strong verb (sub)classes is presented and divided by search code, thus acknowledging the impact of every search code on the identification of inflectional forms for each sub-class, Second, a review of the different inflectional endings, their spelling variants and their distribution along the verb classes is presented. The third part of the results concentrated on the study on complex verbs and the forms adopted by the different pre-verbal items found. Given that Old English is characterised by a strong inflectional system and the use of language internal word-formation processes, the study of derivational elements is justified. Finally, the lemmatisation of the corpus is displayed in a list of 1,768 strong verb lemmas sorted in alphabetical order.

5.3 Conclusions

The current section presents the main conclusions that can be drawn from the analysis carried out in the previous chapters. The conclusions go along two lines. In the first place this thesis provides a motivated answer to the question of the limits of automatisation in morphological analysis. In this respect, it is nesessary not only to present the inventory of lemmas of strong verbs but also to distinguish the automatic searches from the manual revision so as to be able to offer a general assessment of this aspect. Secondly, the conclusions also pose the question of the normalisation of

spellings in headword definition. Indeed, the degree and characteristics of the regularisation of the spelling of the headwords clearly determines the results of the analysis given that different inventories of strong verbs can result from different nomalisation principles. These questions are considered from the qualitative and the quantitative points of view.

5.3.1. Qualitative aspects

To begin with, the analytical steps have led to the creation of a list of lemmas which comprises a total of 1,768 items belonging to all seven classes of strong verbs as well as to all the subclasses. This means an increment of 276 lemmatised forms with respect to the original masterlist as displayed in Appendix 1. There are several reasons for such an increase.

On the one hand, the use of automatised search criteria allows results to be organised in alphabetical order, which implies that similar forms appear close to each other. This has a double effect. First, the identification task is not as time consuming as it would be if the corpus was searched manually. Second, and more important, is that variant spellings are more easily identified and compared with nearby words, which leads both to the identification of a higher number of inflectional forms and, in some cases, to the identification of new lemmas.

On the other hand, some methodological decisions also have an effect on the final amount of lemmatised forms. As already introduced in section 3.6.2, the manner in which the question of circularity has been treated regarding pre-verbal items has an impact on the lemmas selected. Because a canonical form cannot be a variant form of another prefix, two lemmas have to be proposed when inflectional forms have been found containing both preverbal item. Thus, a new lemma *ansittan* has been put forward on the basis of words like *ansittað*, whereas the lemma *onsittan*, already present in the original masterlist has been preserved, corresponding to the unlemmatised forms *onsit*, *onsiteð*, *onsitet*, *onsiteþ*, *onsitt*, *onsittað*; *onsittan*, *onsittaþ* and *onsitte*. A second phenomenon that requires some comment regarding the question of prefixes is that, in some occasions, the separation of affixes has led

to the replacement of a lemma for a new one, instead of to the preservation of the original lemma and the addition of a new form. Let us use the same affixes as illustration. In the original masterlist there was a lemma *onswincan* (IIIa), and I have put forward a new lemma *answincan* through the inflectional form *answincað*. However, no inflectional forms have been found with the pre-verbal element *on-*, and, consequently, *onswincan* has been left out of the final list of lemmas.

Figure 1 below presents the distribution of pre-verbal elements as has been determined in this work to avoid circularity and cross-references:

CANONICAL FORM	**ALTERNATIVE SPELLINGS**
ā-	
ǣ-	
æfter-	*æft-, æftyr-, efter-, eftyr-*
ǣg-	
ǣr-	*ær-, ar-, ars-, ēar-*
æt-	*ad-, æd-, at-, ēt-, et-, ot-*
an-	*ǣn-, ānа-, ann*
and-	*iand-, ond-, ont-*
ante-	
arce-	
be-	*b-, bea-, beah-, beo-, bet-, bew-, bi-, bī-, bic-, big-, bio-, bis-*
eall-	*æl-, æle-, æles-, al-, all-, eal-, el-*
ed-	
efen-	*æfen-, efn-, emn-*
eft-	*æft-, oft-*
ell-	*el-*
em-	
for-	*f'-, f'r-, fær-, færn-, far-, feor-, fer-, fern-*
fore-	*foren-, fores-, forn-, fors-*
forþ-	*fort-, ferþ-*
fram-	*frum-, from-, frun-, frym-, frymft-, frymfþ-*
frēa-	
full-	*ful-*
ge-	*cg-, g-, ga-, gæ-, gæn-, gær-, gad-, gan-, gar-, ged-, gen-, gem-, ger-, gi-, gif-, gim-, gy-*

geond-	*giond-, gind-, gio-, gion-, gin-, geon-, gon-, geo-*
healf-	*heal-*
īg-	*ēh-, eīg-, i-*
in-	*ine-, ing-, inn-*
med-	*me-, met-, mi-, mid-, mið-, mod-*
mis-	*miss-, mus-*
ō-	
of-	*æf-, af-, off-*
ofer-	*eofer-, eofor-, ofær-, ofern-, ofor-, of'-, ofyr-, ouer-, ouyr-, ofer-*
on-	
or-	
oþ-	*oeþ-*
sām-	*sam-*
sin-	
sub-	
tō-	
twi-	*twig-, twy-*
þri-	*þry-, þrie-*
þurh-	*þorh-*
un-	
under-	*und-, undern-, ynder-*
ūp-	*upp-*
ūt-	*utt-, vt-*
ūþ-	
wan-	
wiþ-	
wiþer	*wiþere-, wiþyr-*
ymb-	*ym-, ymbe-, emb-, embe-, eme-, imb-*

Figure 1: The set of pre-verbal canonical forms and their variant spellings.

Finally, I shall refer to the third part of any inflected word, the inflectional ending. Throughout this research, the following endings have been identified:

-að; *-aþ* (*-ad*; *-æd*; *-æð*; *-æþ*; *-at*; *-ath*); *-an* (*-æn*); *-ð*; *-þ* (*-d*; *-th*); *-e* (*-eg*; *-cg*; *-ch*; *-cht*; *-gc*; *-h*; *-o*); *-eð*; *-eþ* (*-cg*; *-ch -cht*; *-et*; *-eth*; *-eað*; *-eaþ*; *-ed*; *-eid*; *-gc*; *-h*, *-id*; *-*

ið; *-iþ*; *-ieð*; *-ieþ*; *-it*; *-ith*; *-oð*; *-ud*; *-uþ*; *-yd*; *-yð*; *-yþ*; *-yt*); *-en* (*-in*; *-yn*); *-est* (*-æs*; *-æst*; *-ast*; *-æt*; *-esð*; *-esþ*; *-is*; *-ist*; *-ust*; *-ys*; *-ysð*; *-yst*); *-on* (*-don*; *-onn*); *-st* (*-sð*; *-s*); *-tst* (*-dst*); *-t* (*-tt*).

The inflectional endings above are presented in alphabetical order considering the canonical spelling. Non-canonical spellings are displayed in brackets.

With these considerations in mind, the set of new lemmas put forward in this dissertation are the following:

Class I: (55) *ædrifan, æfterþīnan, ærīsan, ærrīnan, æswīcan, ætrīnan, āfīgan, anbīdan, anbītan, andbīdan, andlīfan, andwlītan, andwrīþan, anhrīnan, befīgan, besīcan, eftrīnan, forgegīnan, forþīnan, geondgīnan, ināginan, incīnan, ingīnan, inþīnan, oferblīðan, ofergīnan, oferslīpan, ofgīnan, ofrinan, ofscīnan, ofslīþan, ofsnīdan, ōgīnan, ongīnan, onhlīdan, onhlīgan, onrīnan, onscītan, onstrīdan, onwrīþan, oþrīnan, oþþīnan, tōlīþan, tōþīnan, þurhgīnan, þurhwīnan, undergīnan, unrīpan, ūtaginan, ūtgīnan, wiþerflītan, wiþflītan, wiþlīþan, ymbgīnan, ymbsnīdan*.

Class II: (22) *ægēotan*, *ānbūgan*, *āngēotan*, *ānlūcan*, *ānscēotan*, *forbrūcan*, *forflēogan*, *forgēotan*, *forlēodan*, *inhrēosan*, *medscēotan*, *oferscēodan*, *ofersēoþan*, *ofrēodan*, *onāhrēosan*, *onlēogan*, *onsūpan*, *tōbrūcan*, *unbecēosan*, *underlēogan*, *underþēotan*, *unscēotan*.

Class IIIa: (34) *ābringan*, *ēspringan*, *ætwindan*, *āginnan*, *anginnan*, *answincan*, *anwinnan*, *becinnan*, *eallginnan*, *eallsinnan*, *forābringan*, *forebringan*, *foresingan*, *forþbringan*, *inbringan*, *incinnan*, *ingeþingan*, *inginnan*, *intingan*, *medginnan*, *ofþingan*, *onbringan*, *ongeanbringan*, *ongeanwinnan*, *onwindan*, *tōgebringan*, *tōginnan*, *undersincan*, *ungeþingan*, *unþingan*, *ūpspringan*, *ūtginnan*, *ūtrinnan*, *ymblinnan*.

Class IIIb: (26) *ērdelfan*, *ætberstan*, *ætdelfan*, *ætiernan*, *ætspurnan*, *ēweorpan*, *andhweorfan*, *angieldan*, *besweorfan*, *eallbeorgan*, *forbeorgan*, *foriernan*, *framāceorfan*, *inbiernan*, *ofceorfan*, *onbeiernan*, *onceorran*, *ongeaniernan*,

ongeiernan, tōgeteldan, unforspurnan, unweorþan, ūpweorpan, ūtāweorpan, ūtweorpan, ymbceorran.

Class IIIc: (5) *ætbregdan, forebregdan, infrignan, onfrignan, ūpbregdan.*

Class IV: (14) *ætberan, forþbrecan, fullgelan, oferfelan, offelan, ofteran, onbrecan, ongeancuman, ongelan, onhreran, tōbecuman, unfelan, ūtniman, wiþerbrecan.*

Class V: (34) *ǣfetan, ǣgietan, ǣmetan, æretan, ætbiddan, ætfetan, andetan, andgietan, angietan, ansittan, eallgiefan, eftcweþan, inlicgan, ofbiddan, oferbiddan, ofercweþan, ofergesittan, ofgietan, onāsittan, ongeanbiddan, ongeancweþan, ongeansittan, ongesittan, ormetan, orwegan, tōgecweþan, tōgesittan, tōgiefan, tosprecan, ūprecan, ūpsprecan, ūpwegan, wiþangietan.*

Class VI: (28) *ætwadan, analan, anhebban, anwacan, anwadan, forþsceacan, fullwadan, inafaran, inhebban, inhlædan, instæppan, medwadan, medfaran, onalan, onfaran, onsceacan, orhlædan, oþwadan , þurhforfaran, unāhebban, unbehebban, unfaran, ungescieppan, unhebban, untōsceacan, wiþsceafan, ymbfaran.*

Class VII: (1) *tōslǣpan.*

Class VIIa: (11) *andlācan, anscēadan, oflācan, ofsceadan, ongescēadan, onlācan, onspatan, þurhscēadan, ungehātan, unscēadan, ūtāscēadan.*

Class VIIb: (14) *æfterēadan, æthlēapan, anēadan, forēadan, foreēadan, ingangan, inegangan, medēadan, ofgangan, ofergangan, oðgangan, togangan, þurhgangan, ymbgangan.*

Class VIIc: (20) *ætfēallan, æthēaldan, andwēalcan, anhēaldan, anstēaldan, anwēaldan, eftbehēaldan, forefēallan, fullhēaldan, inweaxan, onāhealdan, onfēaldan, onfēallan, ongewēaldan, onstēaldan, sinewēaltan, sinwēaltan, ungewēaldan, ūphēaldan, wiþhēaldan.*

Class VIId: (5) *forbegangan, fullgangan, ofgangan, onspannan, ūpgangan.*

Class VIIe: (21)*ancnāwan, andrædan, anlætan, anrædan, edrædan, eftrædan, forārædan, inswāpan, mislætan, ōcnāwan, ofrædan, onbelætan, orcnāwan, uncnāwan, unlætan, unrædan, ūpārædan, ūtblāwan, ūtlætan, wiþerrædan, ymbsāwan.*

Class VIIf: (7) *edblōwan, healfhlōwan, inblōwan, ofrōwan, onāblōwan, ongeanhlōwan, ōwēpan.*

Class VIIg: (2) *inbūan, onbūan*

All the new lemmas correspond to derived versions of a limited number of verbs. In class VIId, all the instances correspond to complex verbs based on *spannan* or *gangan,* whereas in class VIIg all cases are derivations of *būan*. This tendency is also exemplified by other subclasses although there might be more than two basic verbs. Consider the case of class IV, where the verbs involved are only *beran, brecan, gelan, felan, teran, cuman, hreran and niman,* that is, 8 basic verbs for a total of 14 prefixed lemmas.

As has been shown in Chapter 4, not all these lemmas enjoy the same status. Whereas some of the forms are undisputed, some others should be taken with caution. Among the undisputed ones, two groups can be distinguished. Those which are the result of the application of the system of searches and the methodological guides of this work and which are not found in any other source, and those which have been identified in the DOE during the checking stage. This second group of verbs should have been included in the original masterlist. Verbs belonging in the first group are *ǣspringan* (*æspringæþ*, *æsprungon*), *ansittan* (*ansittaþ*) and *tōbecuman* (*tobecome, tobecumað*, *tobecymð*), whereas the set of verbs in the second group are *ābringan, ætberan, ætberstan, ætbregdan, ætfēallan, æthēaldan, æthlēapan, ætspurnan, ætwindan, āweorpan, āginnan, edblōwan eftbehēaldan,*

foresingan, forflēogan, forþbrecan, forþbrecan, forþbringan, forþsceacan, framāceorfan, fullgangan, ofercweþan and *ofergīnan.*

Consider the infinitives listed below with their corresponding attestations:

I:

oþþīnan (*oððan*).

II:

forlēodan (*forliet*), *onlēogan* (*onlæg*), *underlēogan* (*underlæg*)

IIIa:

incinnan (*incan, inccan, ingcan*), *undersincan* (*undersanc*), *ymblinnan* (*ymlan*).

IIIb:

unweorþan (*unwurð*).

IV:

oferfelan (*oferfel), offelan (offele), ofteran (ofter), unfelan* (*unfæle, ungefel, ungefelan*).

V:

ætbiddan (*ætbist*), *andgietan (andgyt, andgite, andgyte, ondgit, andgiet, ondgiet, andget, andgiete, ondgite, ondgiete*), *angietan (andgytt, andgitt, andgeat*; *angyt, angite, angit), ofbiddan* (*ofgebæd*), *oferbiddan* (*oferbit*), *ongesittan* (*ongesett, ongesittan*), *ormetan* (*ormæte, ormete, ormæt, ormæten, ormeten, ormæton, ormætts), orwegan* (*orweg*), *tōgecweþan* (*tocwædon, togecwæde), ūpwegan* (*upwegað*), *wiþangietan* (*wiðangate*)

VI:

ætwadan (*ætwat*), *ymbfaran* (*emfare, emfaran), inhlædan* (*inhlet*), *medwadan* (*medwad*), *medfaran* (*midfare*).

VIIa: *ongescēadan* (*ongesceat), onlācan* (*onlece*).

VIIc:

onāhealdan (*onahyldað*), *onstēaldan* (*onstealde*, *onstealdan*, *onstealdest*).

VIIf:

owēpan (*owope*).

In all the instances above, the inflectional forms justify the proposed lemmas, and only of those lemmas. However, the data are less conclusive when the inflected forms could belong in two different paradigms. Under these circumstances, the following lemmas have been proposed:

I:

ōgīnan (*ogan*), *ongīnan* (ongan, onginð, ongin, onginst, onginþ, onginen), *þurhgīnan* (þurhgan, ðurhgan), *undergīnan* (undergan), *ūtaginan* (utagan), *ūtgīnan* (utgan), *ymbgīnan* (ymbgan).

II:

forbrūcan (forbrycð, forbricð, forbricþ), *tōbrūcan* (tobrycð, tobrycþ).

IIIa:

anginnan (angunnon), *forābringan* (forabreng, forabrengað), *forebringan* (forebrengað), *inbringan* (inbringan, inbringst, inbringað), *inginnan* (*ingan*, *ingann*, *ingunnen*, *ingunnon*), *medginnan* (*modgan*, *megan*) *onbringan* (onbrincg), *tōgebringan* (togebringð), *tōginnan* (togann), *ūtginnan* (utgan).

IIIb:

onceorran (oncerreð), *ymbceorran* (ymbcerr, ymbcerreð).

IV:

wiþerbrecan (wiðerbrecan, wiðyrbrecan, wiþerbrecan, wiðerbrecaþ).

V:

ofergesittan (ofergesett), *onāsittan*, (onasett), *ongeansittan* (ongeansett), *tōgesittan* (togesett).

VI:

onalan (onælð; onælþ), *oþwadan* (oðwat), *þurhforfaran* (þurhforfærð, þurhforþfærþ).

VIIb

inēadan (ineod, ineode, ineodon, inneode, inneodon, inneade, ineade, ingeode, ineoden, inneadest), *ingeeadan* (ingeodon, ingeeode), *ofēadan* (ofeode, ofeade, ofeodon), *medēadan*, (meteode), *oferēadan* (ofereode, ofereodon, ofereade), *oþēadan* (oðeodon, oþeode) , *tōēadan* (toeodon, toeode), *þurheodan* (þurheode); *ymbēadan* (ymbeode, ymbeade, ymbeodon).

VIIc:

fullhēaldan (fullhealden), *ūphēaldan* (upheald).

VIIe:

orcnāwan (orcnawe).

VIIf:

Inblōwan (inbleow), *onāblōwan* (onableow).

This set of verbs is relevant to the question of the limits of automatic lemmatisation and how the methodological criteria that guide the research process have direct consequences on the form and number of the lemmas that are proposed. Different phenomena can be considered at this point.

In the first place, the choice of root vowel. Given the evolution of the vocalic stems of the Old English strong verbs, the selection of the stem vowel is likely to result in a different set of verbs. Thus, the class III verbs derived from *ceorran* rather than from the weak verb *cyrran*, which would leave them out of the analysis. In the same vein, keeping the *i* of the verbs based on *sittan* causes them to be

included in the inventory of strong verb lemmas, whereas choosing a more modern spelling, such as *settan,* would make us consider them weak verbs.

Choosing a different vowel does not necessarily lead to the change from strong into weak. The inflected forms *inbleow* and *onableow* could be part of the paradigms of *inblāwan* or *inblōwan* and *onāblāwan* or *onablōwan* respectively. In such cases, both lemmas have been put forward.

The case of derived verbs based on *ēodan* is similar. This corresponds to the old form of 'to go', which evolves into the irregular *gangan* and the contracted *gan*. Given that irregular verbs fall outside the scope of this work, *gan* and its derivatives have not be proposed, but complex verbs both based on *ēadan* and *gangan* have been put forward, thus the pairs *inēadan-ingangan*, *oferēadan*-ofergangan and *tōēadan-tōgangan.*

5.3.2 Quantitative matters

Let us begin this section by summarising the figures of inflectional forms identified during the research. The results are offered by class, with the data corresponding to all the subclasses subsumed under classes III and VII.

	Long1	**%**	**Long2**	**%**	**Long3**	**%**	**Filter1**	**%**	**Filter2**	**%**	**Filter3**	**%**	**Total**
I	359	30.12	150	12.61	582	48.51	60	5.06	42	3.69			1,190
II	284	39.08	134	15.90	263	34.95	12	1.88	75	25.60	9	2.56	774
III	555	26.19	206	10.64	775	39.56	290	14.25	175	8.88	10	0.46	1,926
IV	116	19.22	60	10.05	287	40.87	69	13.81	44	6.03			688
V	314	22.03	145	10.23	747	52.16	123	8.78	68	4.77	29	2.00	1,428
VI	215	24.25	68	7.73	532	59.85	35	3.70	37	4.44			883
VII	498	24.35	252	12.29	1,136	55.49	77	3.75	69	3.36	11	0.73	1827
												Total	**8,716**

Table 1. Number of identified inflectional forms by search command.

As can be seen in Table 1, the most frequent verb classes are III, V, and I, with 1,926, 1,428 and 1,190 inflectional forms respectively.

As expected from the manner in which the search commands were designed and implemented, the Long3 search turns out the highest number of hits in all the

classes. Only class II appears as an exception as it is the Long search I that offers the highest number of results. In percentual terms the Long3 search provides around 50% of the results in all classes, with class II showing the lowest percentage of only 34.95% and class VI offering the highest percentage, 59.85%.

The filters, on their part, although responsible for the identification of a remarkable number of inflected forms do not present results comparable to those of the longitudinal searches, especially Filter3. Be it because the previous search codes cover almost fully the range of possibilities or because the specialisation of the filter commands restrict the searches too much, it would seem adequate to review them in future research.

Nonetheless, the first function of the filters is not to contribute to the identification of inflectional forms on their own, but rather, to restrict the otherwise unlimited Long4 search and, in this respect, they have proved to be a more than adequate tool, for they have reduced the results of the automatised search to a maneageable amount of data compatible with manual work. Table 2 summarises the data regarding the reduction of the retrieved results and renders the reduction in percentual terms.

	Long4	**Filter1**	**% Reduction**	**Filter2**	**% Reduction**	**Filter3**	**% Reduction**
I	20,925	2,592	87.32	355	98.3	36	99.83
II	19,550	2,255	88.5	825	95.78	29	99.86
IIIa	7,693	1,009	86.89	168	97.82	0	100
IIIb	9,056	1,043	88.49	256	97.18	14	99.85
IIIc	801	142	82.28	42	94.76	0	100
IV	19,724	1,689	91.44	391	98.02	53	99.91
V	40,079	4,240	89.43	1,048	97.39	53	99.87
VI	26,925	3,085	91.65	794	97.06	20	99.993
VII	134	25	81.35	6	96.53	0	100
VIIa	2,418	325	85.56	62	97.44	0	100
VIIb	3,703	297	91.98	117	96.85	2	99.95
VIIc	2,302	347	84.93	104	95.49	24	98.96
VIId	850	125	85.3	16	98.12	0	100
VIIe	7,092	794	88.81	283	96.01	0	100
VIIf	3,306	356	98.31	86	97.4	0	100

	Long4	Filter1	% Reduction	Filter2	% Reduction	Filter3	% Reduction
VIIg	5,275	605	88.54	263	96.91	10	99.99

Table 2. Percentage of reduction of results by filter.

Table 2 can be interpreted as follows. Long4 puts forward the number of hits obtained after launching that search command. With the exception of classes IIIa, VII, VIIa, VIIc and VIId, all the Long4 searches throw over 3,000 hits, reaching a peak of 40,079 hits for class V. Those data make it impossible to conduct manual work. The use of the filters proves useful as it allows me to carry out the non-automatised task of this project. It is interesting to note how, despite the different figures present in the Long 4 column, the percentages of reduction for each filter is nearly constant throughout the different verb classes. Thus, Filter1 offers reduction percentages within the range of 85-92% with only class IIIc being under it at 82.28% and class VIIf being over at 98.31%. Filter2 also offers a stable range of 94-98% with no remarkable exceptions, whereas Filter 3 offers results over 98%. Therefore, the filters offer a sustained reduction which contributes to the stepwise manner in which this research needed to be carried out.

Along with the ablaut changes and the choice of the vocalic form for the lemmatised forms, a second aspect that presents a great deal of variation is the spelling form of the inflectional ending. In section 5.3.1 above, I provided the list of forms that have been attested. I shall now turn to the quantitative side of the question and describe their distribution among the different verb classes. Table 3 below offers a quantification of the attestations of each inflectional form (both canonical and non-canonical) by verb class. Canonical forms are given in bold type. The symbol ø is used to mark those forms that have no explicit ending.

	I	**II**	**IIIA**	**IIIB**	**IIIC**	**IV**	**V**	**VI**	**VII**	**VIIA**	**VIIB**	**VIIC**	**VIID**	**VIIE**	**VIIF**	**VIIG**
-að	84	44	93	104	16	64	104	80	3	10	10	36	13	24	10	2
-aþ	52	19	36	47	4	15	56	36	1	5	2	18	4	22	4	
-ad	7	2		6		2	19	4		2	1			1		
-æd												2				
-æð	9		10	6		7	16	7		1		3	2	6	2	
-æþ	9	1	8	4		3	13	6		1	2	3		2	2	
-æt	1							1				1				
-at	3	1	2	1		1	6	3		1		1		1		
-ath	2															
-an	184	62	79	102	23	74	122	89	3	11	22	57	21	42	22	1
-æn	7	1	7	9		3	14	2		3		4		3	1	
-e	176	98	115	168	37	156	161	147	8	30	71	99	45	84	32	17
-eg			1													
-o			1			4										
-eð	70	10	48	91	8	49	95	56	2	12	9	25	10	31	10	5
-eþ	38	1	24	25	10	26	55	27		6	4	21	6	20	8	
-cht			1													
-d	4		2		3	4	4	12		2	1			6		
-ð	45	25	64	52	1	50	53	21	1	6	7			15	18	
-et	3		1	2			2	4		1		1		1		
-eth			1			1	1									

	I	**II**	**IIIA**	**IIIB**	**IIIC**	**IV**	**V**	**VI**	**VII**	**VIIA**	**VIIB**	**VIIC**	**VIID**	**VIIE**	**VIIF**	**VIIG**
-eað							6									
-eaþ							3									
-eid							1									
-eth			1			1	1									
-id	1			2		2	1									
-ið	3		2			2										
-iþ			1							1	1					
-ieþ														1		
-it	2						1	1								
-ith			1			1	1									
-oð	1		2	1										1		
-t	25	4	4	1			93	9	1	7		6		11		4
-th						3										
-tt					1											
-þ	14	4	18	9		15	28	21	1					6	7	
-ud				2				1		1						
-uþ	1															
-yd							1	2				1				
-yð	12		5	7	3	4	9	6		1		5		5	2	
-yþ														2		
-en		31	20	48	10	35	177	119	6	18	24	64	16	74	25	

	I	**II**	**IIIA**	**IIIB**	**IIIC**	**IV**	**V**	**VI**	**VII**	**VIIA**	**VIIB**	**VIIC**	**VIID**	**VIIE**	**VIIF**	**VIIG**
-in														1		
-yn	3	4	2	2	1	1	2	3				1		2	2	
-est	16	5	1	10	4	9	44	9		6	4	18	2	22	2	
-æs	1	2	2				6							2		
-æst	2			1	1		2							1		
-ast	1	2	1	3	1	2	2			1		1				
-esð							1									
-esþ			1													
-is						1										
-ist				2	2											
-s				1			11					4		3		
-sð	11						5									
-st	26	10	11	9	10	27	49	9	1	7	2	25		31	20	5
-ust				2	2			1								
-ys						1	1			1						
-yst		2		3	3	1	3	1				3		2		
-on	138	67	84	97	19	64	82	53		7	25	22	4	25	13	9
-onn							1									
ø	277	304	212	224	51	174	326	183	5	76	50	134	46	107	38	19
Total	1245	794	861	1041	205	789	1589	944	32	217	240	551	167	555	217	62

Table 3: Number of instances of each inflectional ending by strong verb class.

The information in Table 3 is based on inflectional ending rather than on person. This is so because many inflectional forms take the same ending and, without an analysis of the context, it is impossible to determine whether they correspond to a given person or not. As has been remarked in chapter 4, formally ambiguous endings are *-e,* which stands for the first person singular of the present indicative and the present subjunctive, both present and preterit; and *-að/-aþ,* which may represent both the indicative plural of the present and the imperative plural.

The first thing worth noticing is the distribution of canonical and non-canonical forms. Canonical forms present a more widespread distribution in which they are present in most verb classes, not to say in all of them. Forms present in all the verbal classes are *-að, -an, -e* and *-eð*. Those endings which are not found in all the sub-classes are usually missing in sub-class VIIg, which presents the most reduced set of inflectional forms. Exception to this general rule are *-en,* which is not found in class I, as all the instances have proven to be of past participles, and *-on,* which has not been found in class VII. Non-canonical forms, however, present a more scattered distribution, and only the forms *-æð, -ð* and *-yn,* can compete with the canonical forms in this respect.

Figures 2 and 3 summarise the distribution of non-canonical forms along the different verb classes. Figure 2 includes those ending having a larger distribution, whereas Figure 3 depicts the situation of those endings having a more limited presence.

ALTERNATIVE INFLECTIONAL ENDINGS	**VERB CLASSES**
-ad	I, II, IIIa, IIIb, IV, VI, VIIa, VIIb, VIIe, VIIf
-æð	I, IIIa, IIIb, IV, V, VI, vIIa, VIIc, VIId, VIIe, VIIf
-æþ	I, II, IIIa, IIIb, IV, V, VI, VIIa, VIIc, VIIe, VIIf
-at	I, II, IIIa, IIIb, IV, V, VI, VIIa, VIIc, VIIe
-æn	I, II, IIIa, IIIb, IV, V< VI, VIIa, VIIc, VIIe, VIIf
-d	I IIIa, IIIc, IV, V, VI, VIIa, VIIb, VIIe
-o	II, IIIa, IV, V, VI
-et	I, II, IIIa, IIIb, V, VI, VIIa, VIIc, VIIe
-oð	I, IIIa, IIIb, IV, VI, VIIe
-yð	I, IIIa, IIIb, IIIc, IV, V, VI, VIIa, VIIb, VIIc,

	VIIe, VIIf
-yn	I, II, IIIa, IIIb, IIIc, IV, V, VI, VIIc, VIIe, VIIf
-ast	I, IIIa, IIIb, IIIc, IV, V, VIIa, VIIc
-yst	IIIb, IIIc, IV, V, VI, VIIc, VIIe

Figure 2. The most common alternative inflectional endings by class.

ALTERNATIVE INFLECTIONAL ENDINGS	**VERB CLASSES WHERE THEY APPEAR**
-æd	VIIc
-ath	I
-th	IV
-ch	IIIb
-cht	II, IIIa
-eg	II, IIIa
-eað	V
-eaþ	V
-ed	VIIe
-eth	II, IIIa, IV, V
-uþ	I
-yt	V
-in	VIIe
-esð	V
-esþ	IIIa
-ysð	VIIc
-sð	I, V
-onn	V

Figure 3: The least common alternative inflectional endings by class.

But the difference between canonical and non-canonical forms is not just a matter of distribution. From a quantitative perspective, in 6,641 inflectional forms out of the 8,716 identified we find a canonical spelling. Other 2,227 forms show no explicit ending, what leaves us with only 662 inflectional forms displaying a non-canonical spelling. This constitutes just a mere 8% of the total explicitly inflected instances. Thus, the preference for canonical spellings in the final part of the lexemes is clearly demonstrated.

It is interesting to focus on the difference between inflectional endings with *ð* or with *þ*. *ð* is far more common if we consider inflectional endings that present both spellings, as in the case with *-að* and *-aþ*. Table 4 explains this question.

Forms with *ð*	**Nº of results**	**Forms with *þ***	**Nº of results**
-að	697	*-aþ*	321
-æð	69	*-æþ*	51
-ð	386	*-þ*	123
-eð	530	*-eþ*	269
-eað	6	*-eaþ*	3
-ið	8	*-iþ*	3
-ieð	1	*-ieþ*	1
-yð	60	*-yþ*	2
-esð	1	*-esþ*	1
Total	**1758**	**Total**	**774**

Table 4: Number of results for dual inflectional endings ending in -ð, -þ.

Some inflectional endings presenting *-ð* or *-þ* are not present in table 3, as their counterparts did not appear in the set of headwords identified in this analysis. Consider the cases of *-oð*, *-uþ*, *-ysð*, and *-sð*, for which no counterpart has been found.

As can be seen in Table 3, there is a clear preference for the choice of the *-ð*. In fact, if we consider the most common forms, namely *-að*, *-ð* and *-eð*, they stand in a 2:1 ratio to their *-þ* counterparts. In the rest of the cases the ratio is similar, although the number of instances is more limited. One case stands out of the rest, the form *-yð*, which occurs 60 times as opposed to only 2 instances of *-yþ*.

With this review of the most relevant quantitative aspects of the research I shall now turn to present some general conclusions and offer an overview of the pending tasks for future research.

5.4 Final remarks and lines of future research

The main conclusions of this thesis, therefore, are that the lemmatisation task has to be partly automatic and partly manual; and that a principled system of regularisation is necessary as it considers the derivational morpheme, the stem and the inflectional morpheme of the word.

The spelling variation and morphological irregularities of Old English texts, caused by the length of the period and for the lack of a standarised orthography, make it impossible to avoid manual revision. In this respect, more contextual analysis is required in future research that refines the results of this work.

Regarding normalisation, the existence of such principles of normalisation does not prevent the rules of regularisation from determining the outcome of the process to a great extent. In future research fine-grained analysis will be required in order to desambiguate some aspects of normalisation like the overlapping of prefixes like *-an*, *-on-*, and *-un*.

Other aspects that call for further attention in future research are the following.

Regarding the automatisation of the search commands, the codes where created on the basis of stable, predictable inflections of verbs, which only accounted for a small number of well-known phonological changes. Thus, the syncopation and the assimilation of the endings *-est-eð*, -e*þ* into *-tt*, *-ð* or *-þ*, just to mention a change in the inflectional ending. Or the vocalic changes in the stem due to *i*-mutation, as the change of <eo> into <ie> in the second and third person singular in the class II verbs. However, other spelling changes have not been considered, such as the substitution of <c> for <k> or <u> for <w>, which are present, according to the DOE, in some attested spellings of the class IV verb *cuman*. These and other changes should be incorporated into the search programme for a more efficient identification of inflected forms, although the impact of such changes as regards the lemmatisation of forms may not be too dramatic.

In future research, the system of searches should be modified so as to apply it to other lexical categories, in such a way that research on the lemmatisation of at least the other major categories could be conducted.

As mentioned above, the question of form-person association is still pending. Whereas that is not a relevant task for the question of lemmatisation, it would be adequate to relate a particular inflectional form to its inflectional information. In the same vein, the question of identical inflectional forms is still open. Some inflectional forms may correspond to two different paradigms. Without the context, it is impossible to say whether they actually belong to either paradigm or to both of them. Given that the decision here adopted has been to postulate two possible infinitives, this is one of the main tasks to undertake in the near future.

6. BIBLIOGRAPHICAL REFERENCES

Bammesberger, A. 1965. *Deverbative jan-Verba des Altenglischen, vergleichend mit den übrigen altgermanischen Sprachen dargestellt*. München: Ludwig-Maximilians Universität.

Bauer, L. 1983. *English Word-formation*. Cambridge: Cambridge University Press.

Bauer, A. H. 2007. Old English *fæst*: a case of grammaticalisation. *Folia Linguistica Historica* 28/1-2: 27-53.

Bergsten, N. 1911. *A study on compound substantives in English*. Uppsala: Almqvist & Wiksells Boktryckeri.

Birkhan, H. 1985. *Etymologie des Deutschen*. Bern-New York: P. Lang.

Blank, A. 2001. Pathways of lexicalization. In M. Haspelmath, E. König, W. Oesterreicher, and W. Raible, (eds.), *Language Typology and Language Universals*. Handbücher zur Sprach- und Kommunikationswissenschaft, 20.2. Berlin and New York, Walter de Gruyter. 1596-1608.

Borowski, B. 1921. *Zum Nebenakzent beim altenglischen Nominalcompositum*. Saale: Halle.

Brinton, L. 1986. *The Development of English Aspectual Systems* Cambridge: Cambridge University Press.

Brinton, L. and E. Closs Traugott. 2005. *Lexicalization and Language Change*. Cambridge: Cambridge University Press.

Bosworth, J. and T.N. Toller. 1973 (1898). *Anglo-Saxon Dictionary*. Oxford: Oxford University Press.

Bussman, H. 1996 .*Routledge Dictionary of Language and Linguistics*. London-New York: Routledge.

Campbell, A. 1987 (1959). *Old English Grammar*. Oxford: Oxford University Press.

Carr, C.T. 1939. *Nominal Compounds in Germanic*. St. Andrews: St. Andrews University Publications.

Clark Hall, J. R. 1996 (1896). *A Concise Anglo-Saxon Dictionary*. Toronto: University of Toronto Press.

Coleman, J. 2006. Lexicography. In B. Aarts and A. McMahon (eds.), *The Handbook of English Linguistics*. Oxford: Blackwell-Wiley. 581-600.

de la Cruz, J. 1973. Old English Pure Prefixes: Structure and Function. *Linguistics* 145: 47-81.

García García, L. 2012a. Morphological causatives in Old English: the quest for a vanishing formation. *Transactions of the Philological Society* 110/1: 112-148.

García García. L. 2012b. Old English *jan*-causatives: between grammar and lexicon. In J. Martín Arista, R. Torre Alonso, A. Canga Alonso and I. Medina Barco (eds.), *Convergent Approaches on Mediaeval English Language and Literature*. Newcastle: Cambridge Scholars. 13-28.

Gardner, T. J. 1968. *Semantic Patterns in Old English Substantival Compounds*. PhD dissertation. Heidelberg Universität.

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. 2010a. The Continuum Inflection-Derivation and the Old English *suffixes -a, -e, -o, -u*. *ATLANTIS* 32/1: 103-122.

González Torres, E. 2010b. The bases of derivation of Old English affixed nouns: status and category. *Studia Anglica Posnaniensia*. 46/2: 21-43.

González Torres, E. 2011. Morphological complexity, recursiveness and templates in the formation of Old English nouns. *Estudios Ingleses de la Universidad Complutense* 19: 45-70.

Guarddon Anelo, M. C. 2009. The Natural Semantic Metalanguage of Old English Compound Adpositions. *ES-English Studies* 30: 61-83.

Hallander, L. G. 1966. *Old English Verbs in* -Sian. *A Semantic Derivational Study*. Stockholm: Almquist and Wiksell.

Haselow, A. 2011. *Typological Changes in the Lexicon. Analytic Tendencies in English Noun Formation.* Berlin: Mouton de Gruyter.

Haspelmath, M. 2002. *Understanding Morphology*. London: Arnold.

Healey, A. *et al. 2004. The Dictionary of Old English in Electronic Form A-G*. 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.

Heine, B. 2003. Grammaticalization. In B. D. Joseph & R. D. Janda, (eds.), *The handbook of historical linguistics.* Oxford: Blackwell. 575-601.

Hendrickson, J. 1948. Old English Prepositional Compounds in Relationship to their Latin Originals. Reprint of *Language Dissertation* No. 43 of the Linguistics Society of America. Philadelphia: University of Pennsylvania.

Hiltunen, R. 1983. *The Decline of the Prefixes and the Beginnings of the English Phrasal Verb*. Turku: Turkun Yliopisto.

Hinderling, R. 1967. *Studien zu den starken Verbalabstrakten des Germanischen.* Berlin: Walter de Gruyter.

Hogg, R., ed. 1992. *The Cambridge History of the English Language I. The Beginnings to 1066*. Cambridge: Cambridge University Press.

Hogg, R. M. and R. D. Faulk 1992 (2011). *A Grammar of Old English*. Oxford, Wiley-Blackwell.

Hohenstein, C. 1912. *Das altengl. Präfix* wið(er)- *im Verlauf der engl. Sprachgeschichte mit Berücksichtigung der andern germ. Dialekte.* Inauguraldissertation zur Erlangung der Doktorwürde der hohen philosophischen Fakultät. Kiel: Christian-Albrechts-Universität zu Kiel.

Holthausen, F. 1963. *Altenglisches etymologysches Wörterbuch*. Heidelberg: Carl Winter Universitätsverlag.

Hopper, P. J. and E. C. Traugott 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

Horgan, D. 1980. Patterns of Variation and Interchangability in some Old English Prefixes. *Neuphilologische Mitteilungen* 91: 127-130.

Jensen, J. 1913. *Die I. und II. Ablautsreihe in der ae. Wortbildung*. Kiel: H. Fiencke.

Kastovsky, D. 1968. *Old English Deverbal Substantives Derived by means of a Zero Morpheme*. PhD Dissertation. Tübingen: Eberhard-Karls-Universität.

Kastovsky, D. 1971. The Old English Suffix *-ER(E)*. *Anglia* LXXXIX 3: 285-325.

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*. Berlin: Mouton de Gruyter. 221-261.

Kastovsky, D. 1989. Morphophonemic alternations and the history of English: Examples from Old English. In M. Markus (ed.) *Historical English. On the occasion of Karl Brunner's 100th birthday*. Insbrucker Beiträge zur Kulturwissenschaft. Anglistische Reihe 1. Innsbruck: Universität. 112-123.

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. 2006. Typological Changes in Derivational Morphology. In A. van Kemenade and B. Los (eds.), *The Handbook of the History of English.* Oxford: Blackwell publishing. 151-177.

Krygier, M. 1994. *The Disintegration of the English Strong Verb System*. Frankfurt am Main: Peter Lang.

Lacalle Palacios, M. 2011. The derivational map of Old English and the limits of gradual derivation. *Revista de Lingüística y Lenguas Aplicadas* 6: 243-259.

Lass. R. 1994. *Old English. A historical linguistic companion*. Cambridge: Cambridge University Press.

Lenze, J. 1909. *Das Präfix* bi- *in der altenglischen Nominal- und Verbal-komposition*. Inauguraldissertation zur Erlangung der Doktorwürde der hohen philosophischen Fakultät. Kiel: Christian-Albrechts-Universität zu Kiel.

Levin, S. R. 1964. A Reclassification of the English Strong Verbs. *Language* 40: 156-161.

Lindemann, J. W. R. 1953. Old English preverbal *ge-:* a re-examination of some current do. *Journal of English and Germanic Philology* 64: 65-83.

Lindemann, J. W. R. 1970. *Old English Preverbal* Ge-: *Its Meaning*. Charlottesville: Virginia University Press.

Lüngen, W. 1911. Das Praefix *ON(D)*- in der altenglischen Verbalkomposition mit einem Anhang über das Praefix. Kiel.

Mailhammer, R. 2006. On the origin of the Germanic strong verb system. *Sprachwissenschaft*. 31/1. 1-52.

Mailhammer, R. 2007. *The Germanic Strong Verbs: Foundations and Development of a New System, Trends in Linguistics. Studies and Monographs 183*. Mouton De Gruyter 978-3-11-019878-2.

Mailhammer, R. 2008. The Typological Significance of Ablaut in the (pre-)history of English. *Clerks, Wives, and Historians: Essays on Medieval English Literature*. Peter Lang Pub Inc 978-3039108350.

Maíz Villalta, G. 2011. Assessing the productivity of the Old English *-læcan*. *Miscelanea. A Journal of English and American Studies* 43: 55-72.

Maíz Villalta, G. 2012. Low Productivity indexes: the Old English verbal suffixes *-ettan* and *-læcan*. *RESLA-Revista Española de Lingüística Aplicada* 25: 117-132.

Marckwardt, A. 1942. The Verbal Suffix *-ettan* in Old English. *Language* 17: 275-281.

Martín Arista, J. 2006. Alternations, relatedness and motivation: Old English *a-*. In P. Guerrero Medina and E. Martínez Jurado (eds.), *Where Grammar Meets Discourse: Functional and Cognitive Perspectives*. Córdoba: Servicio de Publicaciones de la Universidad de Córdoba. 113-132.

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. Adjective formation and lexical layers in Old English. *English Studies* 92/3: 323-334.

Martín Arista, J. 2011b. Morphological relatedness and zero alternation in Old English. In P. Guerrero Medina (ed.), *Morphosyntactic Alternations in English*. Sheffield/Oakville: Equinox. 339-362.

Martín Arista, J. 2012. Nerthus. The reference list of Old English strong verbs. *Working Papers in Early English Lexicology and Lexicography* 2.

Martín Arista, J. 2012. The Old English Prefix *ge-:* A Panchronic Reappraisal. *Australian Journal of Linguistics* 32/4: 411-433.

Martín Arista, J. 2013. Recursivity, derivational depth and the search for Old English lexical primes. *Studia Neophilologica* 85/1: 1-21.

Martín Arista, J. 2013. 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., E. González Torres, G. Maíz Villalta, R. Mateo Mendaza, C. Novo Urraca, R. Vea 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. From directionals to telics: meaning construction, word-formation and grammaticalization in Role and Reference Grammar. In M. A. Gómez González, F. Ruiz de Mendoza Ibáñez and F. Gonzálvez García (eds.), *Form and function in language: Functional, cognitive and applied perspectives*. Amsterdam: John Benjamins. Forthcoming.

Mateo Mendaza, R. 2012. The Old English adjectival suffixes *-cund* and *-isc*: textual occurrences and productivity. *ES-Revista de Filología Inglesa* 33. 197-213.

Mateo Mendaza, R. 2014. The Old English adjectival affixes *ful-* and *–ful*: a text-based account on productivity. *NOWELE-North-Western European Language Evolution.* Forthcoming.

McArthur, T. 1992 *The Oxford Companion to the English Language*. Oxford and New York, Oxford University Press.

Meroney, H. 1943. *Old English* upp, uppe, uppan *and* upon. University of Chicago. PhD Dissertation.

Mitchell, B., C. Ball and A. Cameron. 1975. Short titles of Old English texts. *Anglo-Saxon England* 4: 207-221.

Mitchell, B. 1978. Prepositions, adverbs, prepositional adverbs, postpositions, separable prefixes or inseparable prefixes, in Old English? *Neuphilologische Mitteilungen* 79: 240-257.

Nicolai, O. 1907. *Die Bildung des Adverbs im Altenglischen*. Kiel: H. Fiencke.

Niwa, Y. 1966. The prefix *ge-* added to *niman* in the OE Gloss to the Lindisfarne Gospels. *Studies in English Literature* (Tokyo). English number, 1966: 65-79.

Palmgren, C. 1904. *English Gradation-Nouns in Their Relation to Strong Verbs*. Uppsala: University of Uppsala.

Pilch, H. 1953. Das AE. Präverb *ge-*. *Anglia* LXXI: 129-140.

Pilch, H. 1970. *Altenglische Grammatik*. München: Max Hueber Verlag.

Quirk, R. and C. L. Wrenn. 1994 (1955). *An Old English Grammar*. DeKalb, Illinois: Northern Illinois University Press.

Ringe, D. 2006. *From Proto-Indo-European to Proto-Germanic*. Oxford, Oxford University Press.

Roberts, J. 1980. Old English *un-* 'very' and *unferth*. *English Studies* 61: 289-292.

Röhling, M. 1914. *Das Präfix* ofer- *in der altenglishen Verbal- und Nominal-komposition mit Berücksichtingung der übrigen germanischen Dialekte*. Inauguraldissertation zur Erlangung der Doktorwürde der hohen philosophischen Fakultät. Heidelberg: Carl Winters Universitätbuchhandlung.

Samuels, M.L. 1949. The *Ge-* Prefix in the OE Gloss to the *Lindisfarne Gospels*. *Transactions of the Philological Society* 1949: 62-116.

Sapir, E. 1921 *Language: An Introduction to the Study of Speech*. New York: Harcourt.

Schabram, H. 1970. Bemerkungen zu den AE. Nomina Agentis auf *-estre* und *-icge*. *Anglia: Zeitschrift für englische Philologie* 88: 94-98.

Schön, E. 1905. *Die Bildung des Adjectivs im Altenglischen*. Kiel: Verlag von Robert Cordes.

Schrader, W. 1913. *For- und fore-Verbalkomposita im Verlaufe der englischen Sprachgeschichte*. Inaugural-Dissertation zur Erlangung der Doktorwürde der Philosophischen Fakultät der Königlichen Universität Greifswald. Greifswald: Universität Greifswald.

Schuldt, C. 1905. *Die Bildung der schwachen Verba im Altenglischen*. Kiel: Verlag von Robert Cordes.

Seebold, E. 1970. *Vergleichendes und etymologisches Wörterbuch der germanischen starken Verben*. The Hague: Mouton.

Siemerling, O. 1909. Das Präfix *for€-* in der altenglischen Verbal- und Nominalkomposition. Inauguraldissertation zur Erlangung der Doktorwürde

der hohen philosophischen Fakultät. Kiel: Christian-Albrechts-Universität zu Kiel.

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.

Talentino, A. V. 1970. *A Study of Compound* hapax legomena *in Old English Poetry*. PhD dissertation. Binghampton: State University of New York.

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. 2011a. The Morphological Structure of Old English Complex Nouns. *ATLANTIS* 33/1: 127-146.

Torre Alonso, R. 2011b. Affix Combination in Old English Noun Formation: Distribution and Constraints. RESLA-*Revista Española de Linguística Aplicada* 24: 257-279.

Torre Alonso, R. 2012. Recursive suffixation in the formation of Old English nouns. In J. Martín, R. Torre Alonso, A. Canga Alonso and I. Medina Barco (eds.), *Convergent Approaches on Mediaeval English Language and Literature*. Newcastle: Cambridge Scholars. 63-86.

Trips, C. 2009. *Lexical Semantics and Diachronic Morphology. The development of -hood, -dom and -ship in the History of English.* Tübingen: Max Niemeyer Verlag.

von Lindheim, B. 1958. Die Weiblichen Genussufixe. *Anglia: Zeitschrift für englische Philologie* 76: 479-504.

Weick, F. 1911. *Das Aussterben des Präfixes* ge- *im Englischen*. Darmstadt: Wintersche Buchdruckerei.

Weyhe, H. 1910. *Zu den altenglischen Verbalabstrakten auf -nes, und -ing -ung. Habilitationsschrift*. Philosophischen Fakultät der Universität Leipzig. Borna-Leipzig: Buchdruckerei Robert Noske.

Wiltshire, C., and A. Marantz. 2000. Reduplication. Morphology. *An international handbook on inflection and word-formation*. Walter de Gruyter: Berlin-New York. 557-567.

Wodtko, S. D., B. Irslinger and C. Schneider. (2008) Nomina im Indogermanischen Lexikon. *Lingua Posnaniensis*. 51/1: 173-180.

Wrenn, C. L. 1933 “Standard” Old English. *Transactions of Philological Society* 32/1: 65-88.

Wright J. and E. M. Wright. 1984 (1925). *Old English Grammar*. Oxford: Oxford University Press.

Zieleger, D. 2003 Redefining unidirectionality: Insights from demodalisation. *Folia Linguistica Historica* XXIV: 225-266.