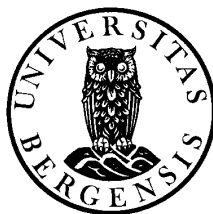


# **A study of empty referential pronominal subjects in Old English**

**Kristian Andersen Rusten**

A thesis submitted for the degree of Master of Philosophy in English  
Department of foreign languages  
University of Bergen  
September 2010





## **ACKNOWLEDGEMENTS**

Thanks are first and foremost due to my supervisor, Kari Elaine Haugland. Her encouragement, advice and, above all, critical reading, have been of great aid in the process of writing this thesis. I also acknowledge my debt to Susan Pintzuk of the University of York, whose kind advice in using the more complex features of the York-Toronto-Helsinki Corpus of Old English Prose and the associated CorpusSearch program has been invaluable. It goes without saying that the blame for any erroneous representation of her advice lies wholly with me. I would also like to thank Gard Buen Jensen for assisting me in the use of the YCOE in the earliest stages of the project.



## TABLE OF CONTENTS

Acknowledgements .....	i
Table of contents .....	iii
List of tables and figures .....	v
Abbreviations .....	vi
<b>1 INTRODUCTION .....</b>	<b>1</b>
1.1 Aim and scope .....	1
1.2 Background .....	1
1.3 Relevance of the thesis .....	4
1.4 Some remarks on Old English .....	5
1.4.1 Note on the poetic tradition .....	5
1.5 The corpus .....	6
1.5.1 Early Old English texts .....	6
1.5.1.1 <i>Bede</i> .....	6
1.5.1.2 <i>Boethius</i> .....	7
1.5.1.3 Manuscript A of the <i>Anglo-Saxon Chronicle</i> .....	7
1.5.1.4 <i>Cura Pastoralis</i> .....	8
1.5.1.5 The <i>Laws</i> of Alfred .....	8
1.5.1.6 <i>Orosius</i> .....	8
1.5.2 Late Old English texts .....	9
1.5.2.1 Byrhtferth's <i>Manual</i> .....	9
1.5.2.2 The <i>Dialogues</i> of Gregory the Great .....	9
1.5.2.3 <i>The West-Saxon Gospels</i> .....	9
1.5.2.4 Ælfric's <i>Catholic Homilies</i> .....	10
1.6 Research questions .....	10
1.7 Structure of the thesis .....	11
<b>2 PREVIOUS RESEARCH .....</b>	<b>12</b>
2.1 Introduction .....	12
2.2 On theoretical affiliation .....	12
2.3 Empty pronouns and the <i>pro</i> -drop parameter .....	13
2.4 Sigurðsson (1993) .....	17
2.5 Pogatscher (1901) .....	18
2.6 Mitchell (1985) .....	20
2.7 Empty pronominal subjects in other early Germanic languages .....	22
2.7.1 Gothic .....	22
2.7.2 Old High German .....	24
2.8 The status of the research tradition .....	26
2.9 Summary .....	27
<b>3 METHOD .....</b>	<b>28</b>
3.1 Introduction .....	28
3.2 Note on terminology .....	28
3.3 Limiting the scope .....	29
3.4 Corpus Linguistics .....	31
3.4.1 The York-Toronto-Helsinki Corpus of Old English Prose .....	32
3.4.2 Data collection .....	34

3.4.3 Problems in using the YCOE corpus .....	37
3.5 Method of analysis .....	39
3.5.1 Problems in using FileMaker Pro.....	40
3.6 Problems of analysis .....	41
3.6.1 Are the selected texts representative? .....	41
3.6.2 Could $S_{\emptyset}$ be a result of scribal error? .....	44
3.7 Summary .....	45
<b>4 DATA AND ANALYSIS .....</b>	<b>46</b>
4.1 Introduction.....	46
4.2 The relative distribution of $S_{\emptyset}$ in OE.....	46
4.3 The historical development of $S_{\emptyset}$ .....	48
4.4 Introductory remarks on the analysis and categorisation of the instances of $S_{\emptyset}$ .....	50
4.5 Hortative subjunctives .....	51
4.6 <i>Hatan</i> + empty subject relative pronouns .....	57
4.7 Other uses of $S_{\emptyset}$ .....	59
4.7.1 The clausal distribution of $S_{\emptyset}$ .....	60
4.7.1.1 $S_{\emptyset}$ in main clauses.....	61
4.7.1.2 $S_{\emptyset}$ in second conjuncts .....	63
4.7.1.3 $S_{\emptyset}$ in relative clauses.....	65
4.7.1.4 $S_{\emptyset}$ in adverbial clauses .....	66
4.7.1.5 $S_{\emptyset}$ in <i>þæt</i> -clauses.....	67
4.7.1.6 Summary .....	69
4.7.2 Syntactic function of the antecedent .....	70
4.7.2.1 $S_{\emptyset}$ co-referent with a preceding or following subject .....	70
4.7.2.2 Oblique experiencer antecedents .....	75
4.7.2.3 $S_{\emptyset}$ co-referent with a preceding object.....	77
4.7.2.4 $S_{\emptyset}$ co-referent with an object of a preposition .....	80
4.7.2.5 $S_{\emptyset}$ co-referent with a genitive .....	81
4.7.2.6 “Miscellaneous” antecedents.....	81
4.7.2.7 “Indefinite” antecedents .....	83
4.7.3 Location of the antecedent .....	84
4.8 “Distance” between $S_{\emptyset}$ and antecedent .....	85
4.9 Identification of the antecedent and possible explanations for $S_{\emptyset}$ .....	87
4.10 Closing discussion .....	96
4.10.1 Predominance in second conjunct clauses .....	96
4.10.2 Subject-antecedent predominance.....	99
4.10.3 Distance/identification .....	99
4.11 Summary .....	101
<b>5 SUMMARY AND CONCLUSION .....</b>	<b>103</b>
<b>REFERENCES .....</b>	<b>106</b>
<b>SUMMARY IN NORWEGIAN .....</b>	<b>110</b>
<b>APPENDIX: Example of a FileMaker record</b>	

## LIST OF TABLES AND FIGURES

Table 2.1	Referential pronouns in OHG: the distribution of overt and empty pronouns in some OHG texts .....	24
Table 4.1	Pronominal subjects in Old English prose: $S_{\text{pron}}$ vs. $S_{\emptyset}$ .....	47
Table 4.2	$S_{\text{pron}}$ vs. $S_{\emptyset}$ in eOE and IOE prose .....	49
Table 4.3	Hortative subjunctives compared to total occurrences of $S_{\emptyset}$ , acc. to texts .....	54
Table 4.4	$S_{\text{pron}}$ vs. $S_{\emptyset}$ with some hortative subjunctive verbs in LawAf.....	56
Table 4.5	The distribution of $S_{\emptyset,\text{rel}}$ according to texts.....	58
Table 4.6	The distribution of $S_{\text{pron}}$ vs. $S_{\emptyset}$ according to clause type.....	61
Table 4.7	$S_{\text{pron}}$ vs. $S_{\emptyset}$ in main clauses according texts .....	62
Table 4.8	$S_{\text{pron}}$ vs. $S_{\emptyset}$ in second conjunct clauses according to texts .....	64
Table 4.9	$S_{\text{pron}}$ vs. $S_{\emptyset}$ in relative clauses according to texts .....	66
Table 4.10	$S_{\text{pron}}$ vs. $S_{\emptyset}$ in adverbial clauses according to texts .....	67
Table 4.11	$S_{\text{pron}}$ vs. $S_{\emptyset}$ in <i>þæt</i> -clauses according to texts.....	68
Table 4.12	Clausal location of the antecedent of $S_{\emptyset}$ .....	85
Table 4.13	The distance between $S_{\emptyset}$ and its antecedent .....	86
Table 4.14	Example of OE verbal inflections: <i>helpan</i> ‘help’ (strong, class 3b verb) .....	88
Table 4.15	Example of OE verbal inflections: <i>hælan</i> ‘heal’ (weak, class 1 verb) .....	88
Table 4.16	Distribution of $S_{\emptyset}$ according to person.....	89
Figure 3.1	An example of the syntactic annotation in YCOE.....	33

## ABBREVIATIONS

### General

A	Accusative case
$\chi^2$	Chi-square
D	Dative case
eOE	Early Old English
G	Genitive case
GB	Government and Binding (or Government-Binding)
IMP	Imperative mood
IOE	Late Old English
NP	Noun Phrase
OE	Old English
OF	Old French
OHG	Old High German
OIce	Old Icelandic
ON	Old Norse
PDE	Present-day English
PL	Plural
POE	Prehistoric Old English
RFLX	Reflexive
S <sub>∅</sub>	Empty pronominal subject
S <sub>pron</sub>	Overt pronominal subject
S <sub>∅.rel</sub>	Empty relative pronoun subject
S <sub>pron.rel</sub>	Overt relative pronoun subject
SBJCT	Subjunctive mood
SG	Singular
V2	Verb-second
YCOE	The York-Toronto-Helsinki Corpus of Old English Prose

### Texts

Bede	The Old English version of Bede's <i>Historia Ecclesiastica Gentis Anglorum</i>
Bo	King Alfred's Old English version of Boethius' <i>Consolatio Philosophiae</i>
ByrM	Byrhtferth's <i>Manual</i>
ChronA	The Anglo-Saxon <i>Chronicle</i> , Ms A (in <i>Two of the Saxon Chronicles Parallel</i> )
CP	<i>Cura Pastoralis</i>
GD	Gregory's <i>Dialogues</i>
LawAf	The <i>Laws</i> of Alfred
Or	The Old English <i>Orosius</i>
WSCp	The West-Saxon <i>Gospels</i>
ÆCHom I	Ælfric's <i>Catholic Homilies</i> , the first series
ÆCHom II	Ælfric's <i>Catholic Homilies</i> , the second series
ÆLS	Ælfric's <i>Lives of Saints</i>



## CHAPTER 1

### INTRODUCTION

#### 1.1 Aim and scope

The present thesis is concerned with the non-expression of referential subject pronouns and the resulting “subjectless” sentences occurring in Old English (OE) prose. The study is based on the 450 occurrences of such sentences in a corpus consisting of a wide selection of Old English texts. The aim of the investigation is to describe the non-expression of referential subject pronouns, and attempt to determine whether any systematicity can be seen with regards to its syntactic – and to some extent – pragmatic “behaviour”. Such an endeavour will primarily necessitate syntactic descriptions of the empty pronominal subjects in evidence, but occasional reference will also be made to certain pragmatic characteristics of these empty subjects. In investigating the status of such empty pronominal subjects, it is hoped that the study can make a valid and useful contribution to the field, particularly in helping concretise an understudied subsection of the syntax of Old English through adding hard data to an area which has traditionally been somewhat impressionistic.

#### 1.2 Background

A typological division can be made between languages which require overt referential subject pronouns and languages which do not. In the various paradigms associated with generative syntax, this phenomenon has been much discussed under the heading of *pro*-drop or *null subject*, based on the assumption that there is an “empty” pronoun present in the clause structure, despite the fact that it is not phonologically or graphically expressed. Languages which do not require overt referential subject pronouns in clauses containing a finite verb have in generative terms been labelled *pro*-drop languages. In languages of this type – archetypically exemplified by Romance languages such as Italian and Spanish – clauses regularly do not feature overt subject pronouns, except for emphasis, and may in this way be considered “subjectless”. Non-*pro*-drop languages, on the other hand, such as English, German and French, require subject pronouns to be overtly present in the clause structure.<sup>1</sup>

---

<sup>1</sup> It may be that “subjectless” clauses are a possibility in regional variants of some of these languages, as is partially the case in Modern German (Axel 2005: 37). Such regional varieties will not be considered here. Also, in agreement with among others Haegeman (1996), “telegram language” – or its newer equivalent, “SMS language” – are not considered genuine examples of the type of *pro*-drop or “subjectless” clauses examined here.

The Modern Germanic languages and French, then, are firmly non-*pro*-drop. As such, these languages represent a divergence from the general rule observed among the world's languages, of which the majority allow *pro*-drop to some extent.<sup>2</sup> There is, however, considerable agreement that the Germanic languages have developed from a stage where overt subjects in fact were not obligatory in the clause structure. As Fertig (2000) points out, Proto-Indo-European, the postulated ancestor language of all the Indo-European languages, is assumed to have been a *pro*-drop language. Evidence from Gothic also suggests that Proto-Germanic – the partially reconstructed ancestor of the Germanic languages – was also a language of this variety (Fertig 2000: 8). If these assumptions are correct, it would be expected that the attested early Germanic texts may reflect this in some manner. Yet, as Fertig also points out, “by the time texts started appearing in the North and West Germanic dialects, null referential subjects had become the exception rather than the rule” (Fertig 2000: 8). This increase in use of overt subject pronouns must be considered an innovation in and specific to the Germanic languages.

While it may be true that empty subjects had become increasingly exceptional by this time, the attested Old Germanic languages demonstrate a situation where empty referential subject pronouns at least to some extent co-occur with the innovative overt variants. In addition to Gothic, which exemplifies East Germanic, empty subject pronouns have a notable presence in, for instance, Old High German (OHG) and Old Norse (ON).

In conjunction with this, it has been claimed that empty referential subject pronouns also were a more or less regular feature of Old English syntax, despite the fact that most OE clause structures feature an overt subject pronoun. For instance, the number of occurrences of “subjectless” clauses in the OE corpus leads Traugott (1992) to somewhat sweepingly state that “[a] grammatical subject is not obligatory in OE” (Traugott 1992: 170). Also, Baker (2007) asserts that “in situations where Modern English uses a pronoun subject, the Old English finite verb can sometimes express the subject all by itself” (Baker 2007: 105). Mitchell (1985) distinguishes three distinct types of “non-expression” of the subject, and also discusses occurrences of “non-repeated” subjects, which we perhaps would call coordinated subjects (Mitchell 1985: 629). He also claims that empty subjects in OE were an “idiomatic” feature of the language (Mitchell 1985: 633).

Now, it is clear that we are dealing here with a phenomenon recognised by scholars of OE, and even from a very early stage, the phenomenon was linked to the ability of an

---

<sup>2</sup> Gilligan (1987) investigates a sample of 100 modern languages, and concludes that only seven require overt pronouns.

inflected finite verb to identify the subject of the clause. It has been claimed by Ohlander (1943–1944: 107, quoted in Mitchell 1985: 628), among others, that an overt subject pronoun in many cases was redundant, as the verbal inflections invariably would supply the necessary information of person, number and gender needed to identify the subject. Mitchell objects to this, on the grounds that the OE verbal inflections were “too ambiguous” to serve this function, even at such an early stage of the language’s history (Mitchell 1985: 628). Visser (1963: 4) agrees with this, saying that “use of the subject pronoun was the rule” due to “extensive formal syncretism [in the verbal morphology]”. Van Gelderen (2000) does not share his view, stating that “[i]n Old English, pro-drop is quite common” as a consequence of “the strength of the verbal person features” (van Gelderen 2000: 121).

Two examples of OE clauses lacking an overt pronominal subject have been provided below:

- (1.1) *Forðæm hit is neah þære tide þe ic getiohhod hæfde on oðer weorc to fonne,*  
because it is near the time which I intended had on other work to do  
*& get næbbe [S<sub>ø</sub>] þis gedon;*  
and yet not-have [I] this done  
‘Because the time is approaching when I had intended to do other work, and yet have I not done this.’  
(Bo 127.26)
- (1.2) *Ac se se ðe unwærlice ðone wuda hiewð, & sua his freond ofsliehð, him bið nidðearf*  
but the-one who unwarily the wood hews, and thus his friend kills, him-D is necessary  
*ðæt he fleo to ðara ðreora burga anre, ðæt [S<sub>ø</sub>] on sumere ðara weorðe genered,*  
that he flee to of-the three cities one-of, that [he] in one of-them become saved,  
*ðæt he mote libban;*  
so-that he may live  
‘But he who hews the wood unwarily, and thus kills his friend, for him it is necessary to flee to one of the three cities, so that he may be saved in one of them, so that he may live.’  
(CP 167.15)

As is evident, empty pronominal subjects of the kind exemplified above are incompatible with the grammatical rules of the modern language, which requires overt pronouns in both cases. Indeed, there are only very few environments in which standard Present-day English (PDE) permits omission of referential subject pronouns, most notably, perhaps in imperative main clauses. Two such clauses have been provided below, one exemplifying PDE and one exemplifying OE.

- (1.3) *Shut the door.*

- (1.4) *Gyf ðu sy godes sunu: cweð to bysum stanum þæt hie beon awende to hlafum.*  
if you be god's son: say to these stones that they become turned to bread  
'If you are the son of God, say to these stones that they should become bread.'  
(ÆCHom I 266.12)

Empty subjects are also commonly observed in second conjunct clauses when co-referent with an overt subject in the immediately preceding "first" conjunct main clause. Again, two examples of such clauses have been provided, illustrating both PDE and OE.

- (1.5) *Peter studied hard and was rewarded at the exam.*

- (1.6) *Ða aras Iosep of swefene & dyde swa Drihtnes engel him behead*  
then rose Joseph from sleep and did as Lord's angel him commanded  
'Then Joseph rose from his sleep, and did as the Lord's angel commanded him.'  
(WSCp Mt 1.24)

As the constructions illustrated in (1.3–1.6) are still idiomatic in the modern language, they do not represent the type of empty subject with which this thesis is concerned.<sup>3</sup>

### 1.3 Relevance of the thesis

While the scholars mentioned above all agree that empty referential subject pronouns occur with some frequency in the language, it seems little consensus has been reached with regard to the actual extent of the phenomenon. This is perhaps reflected most clearly in the widely differing accounts of the role played by verbal inflections in Mitchell (1985) and van Gelderen (2000). The fact is that very little systematic research has been done on empty referential subject pronouns in OE, and although many somewhat categorical claims have been made about their occurrence, the fact remains that very little hard data on the topic has been gathered. While it should be conceded, in fairness, that the tools necessary for efficient study of this area of the language – namely syntactically annotated corpora – have not been available until recent years, the main goal of this study is to remedy this lack of hard data.

As we have seen, then, no in-depth study of the phenomenon has been carried out as of yet, nor have anyone, to the best of my knowledge, endeavoured to thoroughly quantify the various instances of "subjectless" clauses in OE. The present thesis will attempt to begin filling the proverbial gap in these areas, although due to the restrictions inherent to a study on this level, a full account will necessarily be impossible. It is simply beyond the scope of this thesis to shed full light on an understudied area in the time and space allocated. Still, it is

---

<sup>3</sup> More will be said on this in Chapter 3.

hoped that this work – if nothing else – may help concretise a formerly rather vague area, and provide some basis for potential further studies on the subject.

#### **1.4 Some remarks on Old English**

Old English is the language spoken and written in the area that would be known as England *c.*700–*c.* 1100 CE. An undocumented period of the language, *c.*450–*c.*700 is referred to as the Prehistoric Old English (POE) period. The language was brought to the area by West-Germanic soldiers and settlers, who during the fifth century onwards rapidly colonised the Isles while displacing the native Celts. The Germanic settlers consisted of several distinct tribes, among them Angles, Saxons, Jutes and Frisians – a fact which would later be reflected in several distinct OE dialect areas. The language we commonly refer to as OE is for the main part comprised of texts written in the dialect of the West Saxons. A distinction is made between the early (eOE) and late (IOE) periods of Old English. The first period refers to texts written between *c.*850–*c.*950, and the latter to those written between *c.*950–*c.*1010.

Compared to that of other early Germanic languages, the OE text tradition is quite rich. According to Robinson (1992: 143), “Old English is second only to Old Norse in the volume and variety of texts”. The considerable extant corpus consists of both prose and poetry. Among the many genres included in the prose tradition are religious texts, histories, chronicles and legal texts.

It is, perhaps, also interesting to note that OE is not the direct ancestor of PDE, as the modern version of the language is descended from a dialect spoken in the Mercian and Northumbrian regions of the country. This area was under heavy Scandinavian influence, and by the Middle English period (*c.*1050–*c.*1450) displays multiple features that depart from what one might refer to as common OE usage.

##### **1.4.1 Note on the poetic tradition**

While the OE poetic tradition is notable and extensive, featuring distinct works such as *Beowulf*, *The Battle of Maldon*, *The Seafarer* and *The Wanderer*, it will not be treated in this work, where exclusive focus will be placed on data collected from OE prose texts. The reason for this exclusion should be obvious. When studying a syntactic phenomenon, it is necessary to be confident that the clauses under analysis take their form for reasons of syntax only. This is potentially problematic in poetry, because the syntax of poems is regularly different from that of “everyday” language, and it is in many cases highly difficult to determine whether the text of the poem takes its form for syntactic or stylistic reasons. OE poetry was standardised

to a large extent, and relied heavily on alliteration and strict rhythmic schemes (Baker 2007: 120). It is thus quite possible that the restrictions imposed by metric features could have affected the syntax of poems. The syntactic differences between prose and poetry involve, for instance, omission of subjects and objects, omission of prepositions, more flexible word-order, and so forth (Baker 2007: 144–152).

Of course, it is not unproblematic to simply disregard or disqualify the many examples of empty pronominal subjects in OE poetry as poetic licence – because it may very well be that they do represent a genuine syntactic phenomenon. The problem is that we cannot be sure if they are present due to syntactic or metric considerations. The introduction of poetry into the data material of this work would simply present too many methodological problems to be worthwhile in a study of this size.

## **1.5 The corpus**

As stated above, the corpus for this study is based on a total of 450 citations taken from a variety of Old English prose sources. Texts from both the eOE and IOE periods are duly represented. What follows below is a short presentation of each of the corpus texts.

### **1.5.1 Early Old English texts**

The extant texts of the eOE period are primarily those associated with the court of King Alfred of Wessex. In the main, these texts are translations into the vernacular from Latin originals. The notable exception is the *Anglo-Saxon Chronicle*, along with parts of Alfred's *Laws*. In terms of dialect features, the majority of the translated texts are characteristically West Saxon, an understandable consequence of being translated and written in that area. The exception here is *Bede*, which displays some notably Mercian dialect features (Fulk and Cain 2003: 64).

#### **1.5.1.1 *Bede***

The OE version of the venerable Bede's *Historia Ecclesiastica Gentis Anglorum*, 'history of the church of the English people', is contained in four extant manuscripts, namely Mss T, C, O and Ca. As the title would suggest, the work is a historical text detailing, first and foremost, the history of the English church. The translation from the original Latin has traditionally been said to be the work of King Alfred himself, but while the text was translated during his reign, this is no longer believed to be the case. The translation may, however, have been commissioned by the king (Bately 1991: 72) In terms of both form and content, the translation

is considered to be quite close to the original Latin, and as stated above, contains a notable amount of Mercian dialect features.

### **1.5.1.2 *Boethius***

The work commonly referred to simply by the name of its author, *Boethius* (Bo), is the OE translation of the *Consolatio Philosophiae*. The *Consolation* was written by the Italian consul Ancius Manlius Severinus Boethius while he sat imprisoned and awaiting execution. The work takes the form of a conversation between Boethius and the ‘Lady of Philosophy’, somewhat resembling, perhaps, the dialogues between Plato and Socrates. *Boethius* became a seminal work in medieval Europe, and was translated into a variety of languages, including Old High German and OE. The OE translation is attributed to King Alfred himself. Due to the fact that Alfred was no scholar of Latin, however, it was required that the meaning of the original was explained to him before he subsequently dictated the OE version (Clement 1986: 129). Additionally, the king introduced numerous changes to the OE version, so that the result is “a translation only in the broadest sense of the word” (Fulk and Cain 2003: 54–58). The OE *Boethius* is contained in two manuscripts – Ms C and Ms B.

### **1.5.1.3 Manuscript A of the *Anglo-Saxon Chronicle***

It has been claimed that the *Anglo-Saxon Chronicle* may “be regarded as the most important work written in English before the Norman conquest” (Magoun 1935, quoted in Garmonsway 1954: xv). It is certainly one of the few original vernacular histories of medieval North-West Europe. The *Chronicle* takes the form of a series of annals, covering, in its most extensive form, the period from the Roman conquest of Britain, in the year sixty before *Cristes geflæscnesse* ‘the incarnation of Christ’. The annals describing these ancient times are obviously based upon Latin originals, but the vast majority of the *Chronicle* consists of original West Saxon prose, ending in the year 1154. The manner of topics covered in the *Chronicle* is highly varied, ranging from enumerations of the lives and deaths of kings and bishops, to detailed descriptions of battles, wars and natural phenomena.

Seven extant manuscripts of the *Anglo-Saxon Chronicle* have been preserved. The manuscript forming the basis for the analysis in this work is Manuscript A (ChronA), which is normally held to be the oldest surviving version. It was composed at various stages, and is thus excellently suited to reflect the changing nature of OE. This presents a problem in terms of periodic classification, however. Haugland (2007: 24) classifies the entries up to 955 as eOE, and the following entries as IOE. This system will be followed here, and as six of the

seven citations from ChronA discussed in the present work cover the period 796–921, this text must be regarded as eOE in our context. The fact that the final citation covers the year 993 should not change this.

#### **1.5.1.4 *Cura Pastoralis***

*Cura Pastoralis* (CP), or *Pastoral Care*, is the OE translations of Pope Gregory the Great's *Liber Regulae Pastoralis*. The translation is considered to be the work of King Alfred himself, and the first of the so-called “Alfredian” translations. CP is considered a faithful translation in terms of content, yet many of the complex syntactic structures of the Latin version have been simplified (Fulk and Cain 2003: 52). In this way, the translation can also be said to be quite independent of the Latin original. Thematically, the work deals with the responsibilities of the clergy, and is preserved in two manuscripts – namely the Hatton and Cotton manuscripts.

#### **1.5.1.5 The *Laws of Alfred***

Alfred's *Laws* (LawAf) is a collection of laws contained in 77 sections preceded by a prologue in 49 sections which presents translations of the Ten Commandments along with other excerpts from the Bible. While the sections from the Bible are translated from Latin, Alfred's own laws are in genuine vernacular, “provid[ing] us with something very near to the actual language used by the King” (Wrenn 1967: 197). The *Laws* are preserved in the Parker manuscript, the same manuscript as that containing the *Anglo-Saxon Chronicle*.

#### **1.5.1.6 *Orosius***

The Old English version of Paulus Orosius' *Historiarum adversum Paganos Libri Septem*, ‘seven books of history against the pagans’, was – like the OE version of *Bede* – long considered to have been translated by King Alfred. While Alfred is no longer believed to be the translator, it may be that he requested the work to be carried out (Bately 1991: 72, Fulk and Cain 2003: 65). The translator of the OE *Orosius* (Or) is notable for taking a number of liberties with respect to the source material. Many aspects of the original have been changed according to the translator's tastes and knowledge. For instance, the account of the geography of the British Isles, along with Northern and Western Europe, have been changed, possibly due to the fact that the translator knew more about these areas than the original author. The episodes of Ohthere and Wulfstan are also inserted here, representing genuine OE vernacular in a work which is primarily a translation. The OE *Orosius* is preserved in two manuscripts, Ms L and Ms C.



## **1.5.2 Late Old English texts**

If the early Old English period is associated with the court of King Alfred, the late Old English period must surely be dominated by the many works of Ælfric (c.950–c.1010), abbot of Eynsham. While only two of his texts, the two series of his *Catholic Homilies*, are treated here, he was an extremely productive and pedagogically minded writer, and one of the great literary figures in England before the Norman invasion.

### **1.5.2.1 Byrhtferth's *Manual***

The *Manual* (ByrM) is the companion-piece to Byrhtferth's *Computus*, a work of such complexity that a commentary was needed. The text is written partially in OE and partially in Latin, with large stretches of text being considered genuine vernacular – albeit very high-flowing and academic. Byrhtferth's *Manual* is contained in a single manuscript, Ms A.

### **1.5.2.2 The *Dialogues* of Gregory the Great**

Pope Gregory's *Dialogues* (GD) is constructed as a series of dialogues between the Pope himself and a man called Peter, presumably after the model of Plato. The work is in the main concerned with presenting and discussing various miracles, signs and wonders. GD presents a problematic case in terms of periodical placement. The text is a translation of a Latin original, carried out by bishop Werferth at King Alfred's behest (Giles 1858 [1969]: 141), which places the text firmly in the eOE period. However, all surviving manuscripts are from the IOE period, and from a syntactic point of view, it is difficult to determine whether the text preserves eOE features or introduces characteristics of IOE. This difficulty might cause us to treat GD as a separate category, but it is also felt that on the whole, the syntactic features may be closer to IOE than eOE. There are three surviving manuscripts, namely Ms C, Ms H and Ms O.

### **1.5.2.3 The *West-Saxon Gospels***

The *West-Saxon Gospels* are contained in several manuscripts, all from the IOE period. Although the original translation is unknown, it is likely much older. Which Latin text served as the basis for the translation is also unknown. The *West-Saxon Gospels* constitute the oldest full translations of the four gospels, although earlier versions exist in the shape of glosses in two Latin manuscripts, which are referred to as the *Lindisfarne* and *Rushworth Gospels*. Because the words of the Scripture were considered sacred, and therefore nearly “untouchable”, it might be feared that any examples of empty subjects contained in this text

are the result of a “word-for-word” translations from Latin. However, this is not the impression given by the text itself, which is not characterised by being gloss-like rendering of the original. Although some adherence to the original is expected in a translation of the Bible, the *West-Saxon Gospels* are said to be remarkably autonomous (Haugland 2007: 21–22).

#### **1.5.2.4 Ælfric’s *Catholic Homilies***

*Ælfric’s Catholic Homilies* (ÆCHom) is the OE version of the *Sermones Catholici*, both written by Ælfric, abbot of Eynsham (c.950-c.1010). Ælfric was a highly productive writer, whose works constitute much of the extant IOE corpus of text. Considerable parts of his work are didactic texts, many of which are intended to educate in matters of religion and language. In an introductory statement to the second of the two series of the *Catholic Homilies*, the author states that “I Ælfric the monk have turned this book from Latin books into the English tongue, for those men to read who know not Latin” (cf. ÆCHom II 1.29), yet Thorpe (1843: VI) surmises that Ælfric served not as a “mere” translator, but also as a compiler of material taken from several sources. Among others, Ælfric mentions as his influences one Smaragdus and one Haymo, who wrote sermons and commentaries on the Scriptures (Thorpe 1843: VI). Haugland points out that, while all of Ælfric’s texts in English are based on various Latin originals, the abbot added “extensive comments of his own” (Haugland 2007: 16). It is thus more than likely that considerable stretches of texts represent genuine vernacular OE.

The two series of the *Homilies* contain paraphrases and other forms of discourse based upon the Gospels and the lives of various saints. The work is obviously intended for recitation, as Ælfric explains his decision to “set the matter which I have turned into two books” because he believes that “it were less tedious to hear if one book is read in the course of one year, and the other in the following (Thorpe 1846: 3).

*Ælfric’s Catholic Homilies* are preserved in two manuscripts – Ms A, which contains ÆCHom I, and Ms K, which contains the entirety of the *Homilies*.

### **1.6 Research questions**

While it is certainly difficult to formulate explicit hypotheses for a work which is explorative in focus and primarily aimed toward the proverbial charting of unknown territory, some central questions will still need to be asked. Also, as claims made by others about both the distribution and the characteristics of these empty subjects are examined in the course of this study, many hypotheses will be tested implicitly. The central research questions are formulated as follows:

- Research question 1: To what extent do empty pronominal subjects occur in the selected corpus of OE prose texts?
- Research question 2: To what extent is the occurrence of empty pronominal subjects determined by syntactic criteria such as location in the clause, type of clause, antecedent type, location of antecedent, etc.

Implicitly, then, it will be tested whether these pronominal subjects are as frequent as Baker (2007), Mitchell (1985) and van Gelderen (2000) have suggested. Additionally, it will be tested whether the syntactic environments in which the empty subjects occur have any bearing on the permissibility of this phenomenon in OE. Substantial parts of this study will thus take the form of a qualitative analysis of the role of a number of structural variables proposed by previous scholarship. The particular syntactic criteria to be tested will be presented in Chapter 3. An investigation of the syntactic criteria associated with the occurrence of empty subjects is carried out because it is assumed that empty pronominal subjects occur with some degree of systematicity in OE. Included here will also be an investigation of the role played by the verbal morphology in allowing the occurrence of empty pronominal subjects. It will also be tested whether the eOE texts display higher frequencies for empty subjects than the IOE texts. This will be done in the interest of determining whether any development toward obligatory overt pronominal subjects can be said to occur between the two periods of OE.

### **1.7 Structure of the thesis**

The thesis is organized as follows: chapter 2 presents previous research on empty referential pronominal subjects. First, a survey will be given of generative studies of the phenomenon under the *pro*-drop heading. Further, accounts are provided of research into empty pronominal subjects in OE specifically, and also in three other early Germanic languages. Chapter 3 details the data collection process, as well as the method of analysis. Also included in this chapter is a presentation of some of the central methodological problems encountered. Chapter 4 will then provide in-depth presentations, analyses and discussions concerning the empty subjects examined in this work. Finally, Chapter 5 offers a summary and reiterates the conclusions arrived at in this thesis.

## CHAPTER 2

### PREVIOUS RESEARCH

#### 2.1 Introduction

While some of the most important claims concerning empty pronominal subjects in Old English have already been mentioned (cf. section 1.2), the following chapter will present the previous research on this phenomenon in some more detail. In addition to Attention will also be given to the considerable amount of research on empty subjects in general carried out by generative linguists working within the paradigms of Government and Binding (GB) and Principles and Parameters theory. While there exists a great wealth of research into empty subjects in other languages, there is no extensive body of work regarding empty pronominal subjects as the phenomenon occurs specifically in OE. The arguably most important of these works will be presented in sections 2.5 and 2.6.

These sections will, however, be preceded in section 2.3 by a brief presentation of the main points made in generative studies of empty subjects. Generative linguists have carried out substantial research on the phenomenon, which they commonly refer to as being explainable by the *pro*-drop or *null subject* parameter. In fact, the overwhelming majority of studies concerning empty subjects – and possibly empty categories as such – has been conducted by generativists, and their work holds a natural place in any presentation of previous research on this topic. Section 2.4 will present a short survey of Sigurðsson (1993), a generative study of argument-drop in Old Icelandic (OIce). It is believed that his findings in OIce are relevant and comparable to findings pertaining to OE, due to the close relationship between these two languages. Finally, some notes on empty subjects in two other early Germanic languages, Gothic and Old High German, will be presented.

#### 2.2 On theoretical affiliation

While considerable focus is given in this chapter to generative studies, it should be made clear that this thesis is not affiliated with the theoretical frameworks of generative linguistics or Principles and Parameters theory. Nor does it attempt to provide a critique, or even an exhaustive portrayal, of research into the *pro*-drop parameter. Still, the great output these scholars have provided must be duly considered, and many of their ideas, innovations and accomplishments are undoubtedly of great merit. In short, it would have been a major oversight to ignore the generative work on the *pro*-drop parameter, as these investigations

likely form the bulk of the linguistic knowledge of the phenomenon. However, the various paradigms and frameworks of Chomskyan linguistics will have little bearing on the interpretations and analyses presented later in this thesis.

### 2.3 Empty pronouns and the *pro*-drop parameter

While many languages, Present-day English included, generally do not allow empty pronouns in the clause structure, there are certainly very many that do. Romance languages like Spanish, Italian and Portuguese are among these. In such languages, unstressed pronouns serving as arguments in the clause are regularly unexpressed. The existence of such empty pronouns caught the interest of generative linguists in the late nineteen-seventies, and the task of “determining the nature and distribution of phonetically null but syntactically present entities” claimed a great deal of scholarly attention in the following decades (Jaeggli and Safir 1989: 9). The so-called *pro*-drop parameter was conceived as a means of explaining and categorising the existence and behaviour of such empty pronouns functioning as either subject or object.<sup>1</sup>

Principles and Parameters theory operates with a view of natural languages as being comprised of a number of parameters that may or may not be “set”. For instance, in a given language, the word order parameter may be set to V2, which would result in that language having a V2-constraint requiring the finite verb to come in second place in the clause structure. Set parameters are in turn perceived to “activate” related properties, causing such properties to occur in clusters. As Haegeman (1996) states, “[i]f a language has property X, it will also have property Y and property Z” (Haegeman 1996: 19). Conversely, some parameters are considered to be mutually exclusive, meaning that some parameters and their related properties cannot co-occur. Terms such as *null subject* and *pro-drop*, then, are generative terms used somewhat interchangeably to refer to the same phenomenon, namely the omission or non-expression of a pronoun functioning as subject in a finite clause. Such omitted, or empty, pronouns are referred to by for example Huang as being “null analogue[s] of an overt pronoun” (Huang 2007: 51). Generativists classify languages as *pro*-drop or non-*pro*-drop according to whether or not the language in question permits empty pronominal arguments. Thus, in languages which allow such non-expression, the *pro*-drop parameter is said to be set. Oft-cited examples of this variety are, as previously mentioned, languages such

---

<sup>1</sup> Of course, this work focuses on empty subjects, and will therefore refrain from further discussion of the ramifications of the *pro*-drop parameter in conjunction with object omission. Besides, in a study of OE, it would likely be difficult to make a rigorous distinction between instances of object-drop and “regular” instances of verbs interpreted intransitively.

as Italian, Portuguese and Spanish. In contrast, English and French are frequently highlighted as archetypical non-*pro* drop languages (cf. section 1.2).

Now, it should be pointed out here that *pro*-drop languages do not prohibit the expression of subject pronouns. Referential subject pronouns are commonly expressed in these languages for purposes such as contrast or emphasis. According to Haegeman (1996), overt realisations of subject pronouns are generally used when such usage “has some semantic or pragmatic effect”. Using examples from Italian, she also points out that the only cases where an overt pronominal subject is impossible is in expletive functions in combination with “weather verbs” which would necessitate the insertion of the so-called expletive or “empty” *it* in equivalent English clauses (Haegeman 1996: 21-22). It is also well-known that expletive subjects of this variety are a feature of non-*pro*-drop languages. That is, languages with an overt subject restriction demand the syntactic “gap” to be filled.<sup>2</sup>

The capability of *pro*-drop languages for omitting subject pronouns has traditionally been linked with a strong verbal morphology. This correlation is referred to in generative terms as the *identification hypothesis*. A common argument is that a language with a full system of verbal inflections can allow subject pronouns to be omitted and still have sentences remain unambiguous. The reasoning behind this is that the unique inflectional endings of the verb will correctly identify the antecedent of the empty subject – and therefore also the identity of the referent, and ultimately the meaning of the clause. Additionally, when a language contains detailed and distinct inflections for person, number, and in some languages even gender, the subject pronoun may be perceived not only as recoverable, which is the key point to this “agreement-theory”, but also redundant (Huang 2007: 54). Substantial efforts have thus been made by generativists to correlate rich inflectional paradigms with the *pro*-drop parameter, and claims have been made that languages with rich systems of verbal agreement indeed “are often *pro*-drop languages” (Haegeman 1996: 24).

According to Huang (2007: 54), “[c]rucial evidence in support of this agreement-based theory of null subjects” is available in quite a wide variety of the world’s languages. He lists disparate languages such as Bani-Hassan Arabic, Modern Greek, Modern Hebrew, Pashto and Georgian as examples, and elaborates that they all correspond in the fact that agreement-features determine which arguments may be omitted. For instance, he says that Georgian has

---

<sup>2</sup> However, Zimmermann (2009) claims that Old French (OF) – traditionally described as a *pro*-drop language featured overt expletive subject pronouns “throughout its period”, thus differing “fundamentally from other Romance null subject languages” (Zimmermann 2009: 63). Zimmermann also proposes that OF should be interpreted as non-null subject language, yet one that still allowed empty subjects if “specific structural conditions were met” (Zimmermann 2009: 63, 90).

verbal agreement with all types of argument, and correspondingly permits all arguments to be omitted. In addition, Pashto is cited as allowing dropping of both subjects and objects, “but only if they agree with the verb” (Huang 2007: 55). Interestingly, perhaps, Huang also references generative studies indicating that children set the *pro*-drop parameter at approximately the same time as they begin learning and utilising inflections (Huang 2007: 55).

As is evident, a quite convincing argument is made for the correlation of empty subjects and a strong verbal morphology. However, it is clear that despite the allure of the theory, it is insufficient to explain the *pro*-drop occurring in many of the world’s languages. The example of German is relevant in this context. This language has certainly retained a fairly complex verbal morphology, yet it does not sanction *pro*-drop. German is a firm non-*pro*-drop language, and requires an overt subject, whether nominal or pronominal, referential or expletive, in all declarative and interrogative clauses.<sup>3</sup> Huang (2007: 57) points out that German also “encodes a slightly greater range of person and number agreement specifications in the verb than Portuguese”, which is a *pro*-drop language. The same can also be argued for Icelandic. Conversely, there are languages such as Vietnamese, which permit empty subjects but lack an inflectional morphology entirely (Rosén 1998: 6).

Generativists have struggled with explaining this insufficiency of the original *pro*-drop theory for years, mainly through somewhat convoluted theories aimed at ascertaining what exactly it is that “licenses” and “identifies” these empty pronouns. It could very well be said that in many of the cases, the hypotheses on the governing of *pro*-drop have become too specific to satisfy the stated goal of Generative theories of languages, which is to formulate general linguistic principles valid for all languages. To the best of my knowledge, there is no consensus among generativists with regard to this problem, which accordingly seems to remain largely unsolved.

The fact that generative theorists may not have succeeded in devising a general set of principles describing and predicting the rules of empty subjects has little bearing on this thesis, however. As previously stated, the objective of this work is not to use data from OE to contribute to solving this problem, but to make statements about empty referential subject pronouns as they occur in OE specifically. This means that the many revisions and

---

<sup>3</sup> There are some who classify Modern German as a semi-*pro*-drop language “due to the occurrence of impersonal passives with an empty expletive” (Axel 2005: 27). It is believed that the description of Modern German as a non-*pro* drop language has better justification, as the standard variants of the language most certainly do not allow empty referential subject pronouns. As repeated before, these are also the empty subject pronouns with which this thesis is concerned, and also the kind of *pro*-drop chiefly discussed in this section.

amendments to the original *pro*-drop theory to a large extent may be ignored. Still, it should be kept in mind that the situation is complex, and it is far from an uncontested fact that a full system of inflections should equal *pro*-drop capabilities.

That notwithstanding, it does seem likely that verbal inflections should play a role, at least, in allowing empty subject pronouns. Rosén (1998) touches upon this, as she tentatively relates the development of obligatory overt pronouns in Indo-European languages to a concurrent development toward simpler verbal inflectional systems. However, she admits that it is “difficult to formulate this correlation as a falsifiable hypothesis” (Rosén 1998: 102).

Several generative studies have also proposed the notion that inflections are secondary to pragmatic context in the identification of empty pronominal subjects. The empty pronouns are in these cases considered to be “bound by a [discourse] topic”, which means that the empty subjects correlates with topic-prominence (Gundel 1980: 139, 142, 144). The argument is that an overarching pragmatic focus functioning as topic need not be repeated in every clause.

This would serve to explain why non-inflectional languages such as Vietnamese may feature empty pronominal constituents. Languages like Vietnamese and Chinese are claimed to be organised along the lines of topic–comment instead of the subject–predicate structure with which European scholars might be most familiar (Rosén 1998). In this connection, Rosén’s claim that Old Norse was a topic-prominent language, as opposed to subject-prominent, is certainly an interesting one, given the many common denominators between Old Norse and Old English (Rosén 1998: 38).

Incidentally, it has been suggested that it might be wiser to speak about a general *pro*-add instead of *pro*-drop, as the latter label does seem rather centred on the Germanic languages, and does not take into account the fact that empty pronominal subjects actually are the norm for a great many of the world’s languages (cf. section 1.2). Indeed, Gilligan (1987), a study based on a data set of samples from more than 100 languages, shows that the “vast majority of the world’s languages” do not require overt pronominal subjects. Instead, “pronominal arguments of these sentences are optionally lexical, i.e. they may be phonologically null” (Gilligan 1987: 2). As Haugland (2007: 58) points out, then, it certainly yields very little meaning to speak about *pro*-drop for a language which has never employed pronouns to be dropped from the subject position.



## 2.4 Sigurðsson (1993)

While not directly pertaining to OE, Sigurðsson's generative discussion (1993) of argument-drop in Old Icelandic is highly relevant and potentially illuminating for our purposes, as the common denominators between the languages are many. As Germanic languages, OE and OIce shared considerable linguistic similarities in areas such as vocabulary, syntax and morphology. Sigurðsson notes that OIce had extensive *pro*-drop capabilities, allowing the omission of referential subject pronouns, including also "arbitrary" and expletive subjects, as well as of objects of verbs and prepositions. Importantly, to him, OIce also featured topic drop, defined as "missing arguments that do not behave like a pronominal, but like a variable bound by a null-operator" (Sigurðsson 1993: 247). By this, it is presumably meant that the clause-initial preverbal position is analysed as the topic position, in contrast to all the other positions in which the subject could occur – which are analysed as subject positions. If an empty subject occurs in clause-initial preverbal position, then, that empty subject exemplifies topic-drop. On the other hand, if the empty subject occurs in any of the other possible subject positions, the token exemplifies *pro*-drop.

Modern Icelandic still retains the capability for omission of expletive subjects and still has restricted topic drop, but the language has lost "genuine *pro*-drop of both subjects and objects", a development which occurred as late as the 18<sup>th</sup> and 19<sup>th</sup> centuries. Interestingly, this linguistic change did not coincide with a deterioration of verbal inflections, which remain mostly unchanged from OIce (Sigurðsson 1993: 248-249). In order to explain this according to the identification hypothesis, Sigurðsson postulates that agreement features never played any part in the recoverability of subjects in OIce, despite its rich verbal morphology. Instead, many of the occurrences of empty arguments in OIce are claimed to be *null topics* rather than genuine examples of *pro*. A major argument in favour of this position is that the properties of OIce argument-drop seem to behave in a "crucially different" manner compared to the *pro*-drop seen in some stereotypical null-subject languages. For instance, empty pronouns in OIce were unable to initiate discourse, a fact which provides a clear contrast to languages such as Italian and Spanish. Also, an OIce argument regularly could not be dropped unless it served some form of narrative topic function in the discourse. We are thus dealing with a "separate" category of argument-deletion operating in OIce, which according to Sigurðsson does not follow the established rules of the classical parametric hypothesis.

The notion of topic-drop as a determining factor in the non-expression of subjects is in concordance with the conclusions of Gundel (1980), and also with the findings and conclusions of Rosén (1998) pertaining to Vietnamese, which – again – is a language which

does not feature a verbal morphology, yet still regularly omits arguments. While the majority of the empty pronominal “subjects” thus are considered topics by Sigurðsson, he allows that “at least some” empty arguments must be examples of “genuine *pro*”. He maintains that Icelandic agreement was “nonpronominal” and as a result took “no part in identifying *pro*” (Sigurðsson 1993: 277). Likewise, he also argues that empty topics are not identified by agreement features (Sigurðsson 1993: 250). Those empty subjects which do exemplify genuine *pro* were identified by *free discourse indexing*, which is taken to mean that recoverability is somehow “deducible from preceding discourse” (Sigurðsson 1993: 260), yet these are according to Sigurðsson in the minority. The cases of null-topics are also said to be independent of an antecedent, in that they were “possible with or without” one, whereas genuine subject *pro* “always required an NP antecedent in preceding discourse” (Sigurðsson 1993: 252).

The situation regarding the empty subject and its antecedent is potentially quite complex, and Sigurðsson lists several examples of empty subjects which do not have a singular antecedent that completely matches the identity of the empty pronoun. He identifies, for instance, “split antecedents” and “partial antecedents”. “Split antecedents” are interpreted as cases where several singular antecedents in combination constitute a plural empty subject. “Partial antecedents” are taken to include cases where only part of the antecedent is present in the clause structure, i.e. where a singular antecedent corresponds to a plural empty pronoun, with the implication that the empty pronoun must have additional, unmentioned referents. There are also cases of empty subjects lacking antecedents entirely, but these are classified as empty topics, and not subjects. In terms of distribution, Sigurðsson notes that empty subjects could occur in both main and subordinate clauses (Sigurðsson 1993: 262).

## **2.5 Pogatscher (1901)**

Pogatscher (1901) is one of very few existing studies on empty pronominal subjects as they occur specifically in OE. Its main merit is that it provides a quite comprehensive list of examples of empty subjects in OE prose and poetry, as well as discussions of these. While the majority of Pogatscher’s examples are from poetry, and thus not of direct relevance to this work, which for the reasons discussed in section 1.4.1 focuses on OE prose, his findings and conclusions are highly relevant. Among these is the demonstration that subject pronouns could be left unexpressed in most types of clauses, including main clauses, coordinated clauses – whether syndetic or asyndetic – and subordinate clauses.

In a voluminous survey (Pogatscher 1901: 262-275) of empty pronouns occurring in subordinate clauses sorted according to subordinator, it is exemplified how empty subjects could occur in structures introduced by *ær* ‘before’, *gif* ‘if’, *hu* ‘how’, *siððan* ‘since/after’, *hwæt* ‘what’, *hwy* ‘why’, *þæt* ‘that’ and so forth. Thus, numerous examples are given of adverbial, relative and complement clauses containing empty pronominal subjects.

Supported by his many examples of empty subjects in subordinate clauses, Pogatscher argues that “das pronomen im altenglischen nebensatze nicht ausgedrückt zu werden braucht”, ‘the pronoun in OE subordinate clauses need not be expressed’, and that these instances cannot be regarded as examples of scribal error (Pogatscher 1901: 276). He goes on to list corresponding cases of empty subjects in subordinate clauses in other Old Germanic languages, such as Old High German and Old Saxon. He subsequently states that “missing” subjects, whether occurring in main or subordinate clauses are examples of a language feature common to all the West-Germanic languages (Pogatscher 1901: 276–278).

As a consequence of this West-Germanic kinship, the rules for the omission of subject pronouns in Old High German described by Kraus (1894: 88–98) are also valid for OE, according to Pogatscher. Kraus specifies five main categories of empty subjects as they appear in OHG. Drawing heavily on Kraus, Pogatscher presents his material sorted according to these categories. The categories are primarily determined by the type of antecedent – whether nominative or non-nominative, nominal or pronominal – and where the antecedent is found – whether in clauses preceding or following the empty subject, and whether occurring in main, coordinate or subordinate clauses. Category I comprises the cases where the empty subject is co-referent with a nominative element, while the empty subjects in category II are co-referent with a non-nominative element. For both of these categories, Kraus states that the antecedent of the empty pronominal subject may occur in preceding or following clauses which may be superordinate, subordinate or coordinate to the clause containing the omitted pronoun (Kraus 1894: 88–98). This means that the antecedent could occur in all clause types. For category II, it is specified that the antecedent of the empty subject may be a full NP, a personal pronoun or a possessive pronoun. No reason is given by neither Kraus nor Pogatscher why the same should not also be true for category I. Category III comprises the cases of subject omission where the clause containing the antecedent has no grammatical relationship with the clause containing the empty subject, i.e. the two clauses are not conjoined. Category IV is Pogatscher’s own definition, and differs from that of Kraus. This category indicates that some empty subjects may not agree in number with their antecedent. This means that an empty subject inflected for the plural may be co-referent with a singular

antecedent, such as a collective noun, or several singular elements. He also speculates that the converse may be possible, namely that a singular empty subject may be co-referent with a plural antecedent. Category V includes cases of subject omission where there exists no antecedent recoverable from the context.

Pogatscher finally claims that “das subject nich ausgedrückt zu werden braucht” as long as the concept of the subject is clear from the context and in the hearer’s mind throughout the discourse. This way, the hearer would hypothetically be able to interpret “subjectless” sentences in a way that would still allow for the correct identification of the subject. His term for this is *vorschweben*, which in most cases would be loosely translated as ‘having in mind’ or ‘having a vague idea’ (Pogatscher 1901: 286). This corresponds well, incidentally, with the definition of *given information* provided in Chafe (1976), namely “that knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance (Chafe 1976: 30). Also, as Haugland (2007) points out, this bears strong resemblance to the behaviour of topics in topic-prominent languages, and it is recalled that Rosén claims that ON, another early Germanic language, could be typologically classified as such a language (section 2.3). Attention must also be drawn back to Sigurðsson’s claims that topics were a substantial part of the argument-drop capabilities of OIce, which is a subcategory of ON. It is interesting, in the light of these claims and Pogatscher’s speculation, yet again to accentuate the fact that OE and ON are closely related languages, and might have shared this functionality.

## **2.6 Mitchell (1985)**

Mitchell incorporates much of Pogatscher’s work in his discussion of the subject in OE and its potential for being omitted. In accordance with Pogatscher, he states his belief that non-expression of pronoun subjects “which can be supplied from a preceding clause must be accepted as idiomatic OE”, and rejects the notion that the existence of the phenomenon in OE is due to scribal error (Mitchell 1985: 633). He does, however, find it difficult to explain why it occurs “only spasmodically”, and criticises Pogatscher’s idea of the subject “hovering” (presumably Mitchell’s clumsy translation of *vorschweben*) as a notion “impossible to apply in practice”, thus practically dismissing it. Also, he argues that such an explanation fails to account for why subjects seem to be interchangeably overt and empty in “parallel situations” (Mitchell 1985: 633). Mitchell additionally states his fear that “the reason for the spasmodic appearance of the idiom lies buried with its users” (Mitchell 1985: 634).

Mitchell (1985) operates with a basic demarcation between two types of missing subjects, for which he uses the labels *non-repetition* and *non-expression*. The first type corresponds to empty subjects in second conjunct clauses which are co-referent with the subject of the immediately preceding “first” conjunct clause (cf. section 1.2, examples (1.5) and (1.6)). As previously mentioned, this type of empty subject is outside the scope of this work, on account of being an idiomatic feature of both OE and PDE. This view is also held by Mitchell (1985), who does not consider these subjects omitted, but rather unrepeated. He also includes in this definition cases where an NP serves as subject for two clauses while being modified differently in the respective clauses and cases where the same subject is valid for more than one sentence structure, whether simple or conjoined (Mitchell 1985: 628).

The label of *non-expression* is used for instances of missing subjects where the empty pronoun is not identical with the subject of the immediately preceding clause. Mitchell subdivides this category into three distinct groups, which seem largely to be based on a selection of Pogatscher’s sorting criteria. First discussed are those cases of an empty subject occurring in a second conjunct followed by a verb inflected for the plural, which refers back to a subject element *not* in the plural – such as “a collective noun or two singular elements in the first [of the two conjoined clauses]” (Mitchell 1985: 629). He notes here his belief that modern readers more easily become aware of the “gap” when the second conjunct is not introduced by a co-ordinating conjunction. The second category constitutes those cases where the antecedent is an element which is not the subject of its clause – cases where “the subject has to be inferred from a word or group of words not in the nominative” (Mitchell 1985: 630). This category would then be comprised of empty pronouns which share reference with, in the typical case, either the direct or indirect object of a preceding clause, whether immediately or otherwise. The final category consists of empty subjects in main clauses which are co-referent with a subject in a nearby subordinate clause “and vice versa” (Mitchell 1985: 632). According to Mitchell, they are most frequent in “imperative-like” hortative subjunctive constructions (cf. section 4.5).

Mitchell’s many examples have in common with the citations comprising the corpus of this thesis that they are highly diverse. However, as the case also was with Pogatscher (1901), no quantification is done, and nothing is said of the distribution of the phenomenon in OE. His three categories of “unexpressed” subjects form a good basis for further research, yet they do not go far toward explaining the existence of this phenomenon. Aside from the potential of Latin influence, which he acknowledges in a brief discussion of unusual word order in *Bede*, no hypotheses are put forth designed to explain these “subjectless” clauses.

Mitchell, characteristically, states that “[t]here is room for more work here” (Mitchell 1985: 628). This, then, is what this study attempts to supply.

## **2.7 Empty pronominal subjects in other early Germanic languages**

It was mentioned in section 1.2 that the Germanic languages have developed from a stage where overt subjects in fact were not obligatory in the clause structure. It was also said that Proto-Germanic, the postulated ancestor language of the Germanic languages, was a *pro*-drop language, and that overt pronominal subjects are a syntactic innovation that becomes apparent by the time extant North- and West-Germanic texts appear. It was also seen in section 2.4 that OIce featured extensive *pro*-drop capabilities, and sanctioned deletion of a variety of argument-types. This language also featured topic-drop capabilities, causing omission of clause-initial pronominal subjects. It must be assumed that the same is true for ON as a whole, due to the near-identical status of Old Icelandic and Old Norwegian. We have also seen in section 2.5 that Pogatscher, based on evidence from Old Saxon and Old High German, believes empty subjects to be a feature common to all the West Germanic languages. As indicated in Chapter 1, evidence from Gothic has been taken by some to suggest that this claim can be extended to East Germanic as well, as attested Gothic features very widespread use of empty referential pronominal subjects. The following section will provide a brief overview of the status of research into empty subjects in Gothic and Old High German, in order to supplement the information given on OIce in section 2.4.

### **2.7.1 Gothic**

Gothic is the only East Germanic language for which textual evidence exists. It is also the oldest of the attested Germanic languages. The language is chiefly preserved in writing in a 4<sup>th</sup> century translation of the Bible, contained in various fragmented manuscripts dating from the 5<sup>th</sup> and 6<sup>th</sup> centuries. As such, it is clear that any generalisations one might make may be inherently flawed due to the limited amount of material at hand, and must be taken at face value.

As recalled, pronominal subjects are only rarely left unexpressed in OE (Visser 1963: 4). Gothic contrasts with this in featuring extensive omission of pronominal subjects. Indeed, Streitberg (1906) and Abraham (1993) show that subject pronouns actually are empty in the vast majority of the cases (Streitberg 1906, Abraham 1993 in Haugland 2007: 84). The fact that the Gothic translation of the Bible is based upon a Greek source – and therefore represents a translation of a firmly *pro*-drop language original – may very well have

influenced this syntactic feature, and there are also “fairly numerous instances with overt pronouns where the Greek version would have had ZERO” (Haugland 2007: 84). The fact that the translator has supplied overt pronouns in at least some cases may indicate that the Greek usage in some instances contrasted with that of his native language. This, in turn, may indicate that empty pronominal subjects are not present in the Gothic text exclusively as a consequence of a literal, gloss-like rendering, but also because these subjects were considered an idiomatic feature of the language. According to Fertig (2000), there are also several examples of overt subject that are present in the Greek, but which have been realised as empty in the Gothic text, another piece of evidence favouring the notion that Gothic allowed empty pronominal subjects (Fertig 2000: 10). It can thus be concluded with some certainty that at least some of the cases of empty subject pronouns attested in the manuscripts are likely to be representations of genuine Gothic idiom. Two clauses exemplifying empty subjects in Gothic have been provided below (taken from Fertig 2000; the glosses and translations are added here).

(2.1) *swa wairþiþ sunus mans in daga seinamma. aþþan faurþis [S<sub>ø</sub>] skal manag gaþulan*  
 so becomes son of-man in day his. But first [he] shall-3SG much suffer  
 ‘So will the son of man be in his day. But first he must suffer many things.’  
 (Luke 17:24–25)

(2.2) *nio wisseduþ [S<sub>ø</sub>] þatei in þaim attins meinis [S<sub>ø</sub>]skulda wisan?*  
 not knew-2PL [you] that in those-D father-G mine [I] should-1SG be?  
 ‘Did you not know that I must be about the affairs of my father?’  
 (Luke 2:49)

In terms of clausal distribution, empty pronominal subjects occur in all clause types in Gothic. If it is true, as Fertig says, that Gothic “preserves the null-subject property fully intact”, this is to be expected (Fertig 2000: 10). The overt subject pronouns, on the other hand, are concentrated in a more specific syntactic environment, as is noted by Haugland (2007). Based on Held (1903), she notes that “most of the overt pronouns appear in subordinate clauses” (Haugland 2007: 84).

Fertig (2000) rejects the notion that verbal inflections have a function in allowing empty subject pronouns in Gothic. He claims that “there is no connection between” insertion of overt subject pronouns “and the ambiguity of inflectional endings”. Thus, there are several examples of inserted overt pronouns even when “the verbal ending is completely unambiguous”, while on the other hand, there are many cases where verbs featuring ambiguous endings co-occur with empty subjects (Fertig 2000: 11).

Finally, it is perhaps also interesting to point out that this translation is older by several hundred years than the oldest OE texts. In terms of historical development within the Germanic branch of Indo-European, Gothic can then be said to represent a much “earlier” stage than OE. One can only speculate as to the situation of subject omission in OE at this unattested time, yet it might well be surmised that the phenomenon was more widespread in the Prehistoric OE period which is concurrent with the Gothic fragments.

### 2.7.2 Old High German

Old High German (OHG) is a term covering a variety of early German dialects textually attested between c.700–1100, approximately making it a contemporary language to OE. Unlike the scarcity of textual evidence in Gothic, the preserved material from OHG comprises quite a rich literary tradition. This tradition consists mainly of translations, but also works in the vernacular, such as the *Ludwigslied* ‘the Song of Ludwig’. Among the translated material are works such as *Isidor*, *Tatian* and *Otfrid*. These particular works form – along with Notker’s translation of *Boethius* and the *Benediktinerregeln* ‘The Benedictine Rules’ – the basis of a statistical survey performed by Eggenberger (1961) aimed at determining the frequency of empty pronominal subjects in OHG. Paul (1919) and Sonderegger (1974) describe increasing use of overt pronouns as an innovation in OHG compared to earlier stages of Germanic, a view which is certainly supported by the evidence seen in Gothic if this language is taken to represent an earlier stage of the Germanic languages as a whole. Table 2.1 (extracted and compiled by Haugland 2007: 86, on the basis of Eggenberger (1961)) demonstrates the distribution of empty (ZERO) subject pronouns in five OHG texts.

Table 2.1 Referential pronouns in OHG: the distribution of overt and empty pronouns in some OHG texts (Haugland 2006: 86)

	<b>Benediktinerr.</b>		<b>Tatian</b>		<b>Isidor</b>		<b>Otfrid</b>		<b>Notker, Boeth</b>	
	n	%	n	%	n	%	n	%	n	%
<b>Overt</b>	24	11.4	2,614	71.2	146	72.3	4,753	88.8	439	99.1
<b>ZERO</b>	187	88.6	1,055	28.8	56	27.7	597	11.2	4	0.9
<b>Total</b>	211	100.0	3,669	100.0	202	100.0	5,350	100.0	443	100.0

With the exception of *Boethius*, then, which is a text belonging to the late OHG period, we see very high frequencies of empty subject pronouns. The figure of 88.6% for the *Benediktinerregeln* is particularly striking, especially considering that Modern German does not permit deletion of referential subject pronouns. The frequencies of 28.8% and 27.7%



observed in *Tatian* and *Isidor* are also considerable. It is apparent from these frequencies that subject pronouns were quite regularly left empty at this stage of the language’s development. Indeed, Axel (2005: 28) argues that OHG “allowed genuine *pro* drop”, but that – unlike in “canonical null subject languages”, overt and empty subjects are observed to co-occur. She also hypothesises that “[t]he loss of the null-subject property in late OHG was [...] the result of a grammatical competition between null and overt subject pronouns” (Axel 2005: 28). As will be shown in Chapter 4, these frequencies are much higher than those seen in the majority of the OE texts investigated in this work. This might perhaps be an indication that OE appropriated the “innovation” of overt subject pronouns at an earlier stage than OHG. Examples (2.7) and (2.8) illustrate “subjectless” clauses in OHG (both taken from Axel 2005; glosses and translations are added here).

(2.3) *quam [S<sub>ø</sub>] thô In geiste In thaz goteshûs*  
 came-3SG [he] then in spirit in the temple  
 ‘He came then in the Spirit into the temple’  
 (Tatian 89.31)

(2.4) *In dhemu druhtines nemin [S<sub>ø</sub>] archennemes [...] fater*  
 in the Lord’s name [we] recognise [...] father  
 ‘In the name of the Lord we recognize the father.’  
 (Isidor 279)

As the case was with Gothic, however, the question arises whether the empty pronominal subjects present in translations such as *Tatian* may be the result of non-idiomatic or “direct” translations from the source language – in this case Latin. Axel (2005) shows, though, that overt subject pronouns were added with some degree of regularity in cases where it was felt that the Latin usage was incompatible with that of OHG. According to Axel, this substitution normally took place when the OHG clause had verb-final or verb-late word order, which is the word order commonly seen in subordinate clauses. OHG thus normally had overt subjects in subordinate clauses. This is also pointed out by Eggenberger (1961), who shows that the majority of the overt subject pronouns in OHG are found in subordinate clauses. As seen above, the majority of the instances of overt subject pronouns in Gothic also occur in this clause type. This correspondence between Gothic and OHG is highly interesting, and may lead to speculating that overt subject pronouns may first have started appearing in subordinate clauses and from there percolated into main clauses. If so, the question would remain whether the same is the case in OE. This hypothesis will be tested in Chapter 4.

Axel claims that the use of overt subject pronouns in subordinate clauses is due to the particular word order patterns common in clauses of this type, and speculates that “null subjects were [...] dependent on verb-fronting” (Axel 2005: 33). If this is the case, it would entail that empty subjects are restricted to clauses where the verb occurs to the left in the clause structure. This would ordinarily correspond to declarative and interrogative main clauses with verb-first and verb-second word order (Axel 2005: 31). In OHG, as in OE, these types of word order are as a rule not seen in subordinate clauses. She also points out that “subjectless (declarative) verb-first order is relatively infrequent” (Axel 2005: 33).

Axel relates the empty subjects of OHG to verbal morphology features, and shows that it is only in the third person singular and plural that empty subjects are used predominantly as opposed to overt realisations. In the first and second persons, the overt variant is prevalent. However, she is at a loss to explain this distribution, as it cannot be attributed to the strength of the inflections (Axel 2005: 35).

## **2.8 The status of the research tradition**

As the preceding sections will have shown, the tradition of research into empty referential subject pronouns in Old English is very tentative, and there is a real scarcity of studies aimed at investigating this phenomenon. As has been seen, the prominent works treating these “subjectless” constructions has been largely philological in focus, with the advantages and disadvantages such an approach entails. Pogatscher (1901) primarily endeavours to provide a list of examples, but does not elaborate on the distribution of these empty subjects, and he provides no quantification. Instead, the collected examples are used as the basis from which Pogatscher draws his conclusions about the various aspects associated with the phenomenon, including, for instance, the clausal environment in which the empty subject may occur and the possible types of antecedent with which the empty pronoun may share reference. His work leaves it unclear whether the examples he provides represent rare exceptions or more frequently occurring phenomena. Also, if they do represent a selection – as opposed to being “all there is” – no impression is given of what the total figures may be. The fact is that despite the many examples given, we do not really know very much about the distribution of the phenomenon in OE. Also, the majority of his examples are taken from poetry, which, as noted in section 1.4.1, may be problematic due to the syntactic peculiarities often associated with this genre. It was also seen that Mitchell (1985), while an otherwise breathtakingly comprehensive work of great philological value, bases much of its treatment on the topic on Pogatscher (1901). While certainly not paraphrasal in nature, the treatment of empty subjects

in Mitchell (1985) is very short, and offers no real improvement on the aspects of Pogatscher's study criticised above. In short, it is clear that the established tradition of research is weak in terms of modern linguistic criteria. However, all of the works mentioned here have maintained that this phenomenon is "idiomatic" or permissible in the OE period. In light of this, it is felt that it must be a central focus of this thesis to quantify the phenomenon to as large a degree as possible.

## **2.9 Summary**

This chapter has outlined the research tradition in conjunction with empty subjects in OE. It has also provided a brief survey of generative examinations of the *pro*-drop parameter. In section 2.2 it was stated that, while generative research constitutes the majority of inquiries into the subject matter, this work is independent of the generative tradition. Section 2.3 sketched the background of the *pro*-drop parameter, and briefly indicated some of the problems inherent to the theory and studies based upon it. As an extension of the presentation of generative work on empty subjects, a presentation of Sigurðsson's study of argument-drop in Old Icelandic was provided in section 2.4. Sections 2.5 and 2.6 offered overviews of two of the most important studies of subject omission in Old English, carried out by Pogatscher (1901) and Mitchell (1985). Section 2.7 provided insight into the status of the research into empty referential subject pronouns in Gothic and Old High German, thus providing counterpoint and additional information to the sections dedicated to general *pro*-drop, Old Icelandic and OE. Finally, a summary of the status of the research tradition into empty subjects in OE specifically was presented.

## CHAPTER 3

### METHOD

#### 3.1 Introduction

This chapter will be focussed on detailing the method employed in collecting and analysing the citations in the corpus of data forming the basis of this study. It will also discuss some of the main problems encountered in the process. More specifically, section 3.2 provides a note on the terminology used, and section 3.3 defines the scope of the study. Following this, a short presentation of corpus linguistics as a method for linguistic research is given, along with a short overview of the York-Toronto-Helsinki Corpus of Old English prose (YCOE) and the way it has been utilised to search for occurrences of empty subject pronouns. In order to achieve as high a degree of transparency as possible, the exact search parameters used to collect the data have been reproduced below. Further, the method used in analysing and categorising the collected tokens is explained. Finally, some problems of analysis are dealt with in some detail.

#### 3.2 Note on terminology

In Chapter 2, various terms were used interchangeably to describe empty pronominal subjects. One of the most frequently occurring terms for the phenomenon which can be defined as null-realizations of a subject pronoun has been the generative term *pro-drop*. *Null subject* has also been used in conjunction with the generative research tradition. This work has generally preferred the term *empty pronominal subject*, but the term *empty subject* has also appeared, and will continue to do so. The term *empty subject pronoun* appears when emphasis is placed on the pronominal aspect of the empty argument, rather than the argument's status as subject. The phenomenon has also been represented by broader terms, such as *subject deletion* or *subject omission*. Other possible ways of describing the phenomenon have been observed in related literature, without necessarily being reproduced here. These include descriptions such as *non-overt subject* and *non-lexicalised subject*. In the remaining chapters of the thesis, the phenomenon under analysis will in many cases be referred to simply as  $S_\phi$ . This term will be contrasted by its counterpart  $S_{pron}$ , understood as an overtly expressed pronominal subject.

Additionally, it should be mentioned that some of the previously referenced terms are highly “loaded” in a theoretical sense. This is especially true for generative terms such as *pro-drop*, “small *pro*” and *null subject*. These terms are firmly embedded in parametric theory and

may not necessarily correspond completely with all the data and examples discussed in this study. Since this work does not define itself as part of the generative tradition, it was decided that specifically generative terms would be avoided to as great an extent as possible in the investigative parts of this study.

It should perhaps also be clarified that while  $S_\emptyset$  will be the preferred term in many cases, less theory-specific terms such as *subject omission*, *empty subject*, etc. may and will occur on occasion after this point, but mainly for purposes of prose variety.

### 3.3 Limiting the scope

In conjunction with the various types of empty subjects observed in the extant OE texts, three main types have been distinguished by the researchers responsible for compiling and annotating the YCOE. These are empty subjects elided under coordination, empty expletive subjects, and what might loosely be called “other cases”. It is with these “other cases” this study will concern itself. Examples of all three categories have been provided below:

- (3.1) *Pa wæs he þær gestonden lichomlicre untrymnesse & [S<sub>∅</sub>] forðferde:*  
 then was he there attacked bodily unhealthiness-D and [he] passed-away  
 ‘There he was attacked with bodily infirmity and passed away.’  
 (Bede 232.20)
- (3.2) *Rihtlice [S<sub>∅</sub>] is gecweden þæt heora an underfenge fif pund. and oðer twa.*  
 Rightly [it] is said that them-G one received five pounds and other two  
 ‘It is rightly said that one of them received five pounds, and the other two.’  
 (ÆCHom II 320.74)
- (3.3) *Pa beseah he up. and geseah fela engla werod. on micelre beorhtnysse scinende. and*  
 then looked he up and saw great angel-G host on great brightness shining and  
*ðæra halgena sawla wið his fleogende mid unasecgendlicum leohte. and [S<sub>∅</sub>] afligdon*  
 the saints’ souls toward him-G flying with immense light and [they] caused-to-flee  
*ða deoflu him fram.*  
 the devils him from.  
 ‘Then he looked up and saw a great host of angels, shining with great brightness, and  
 the souls of the saints flying toward him with immense light, and they caused the  
 devils to flee from him.’  
 (ÆCHom II 194.155)

Example (3.1) illustrates an empty subject elided under coordination, i.e. a case where the subject of a second conjunct clause has been deleted on account of being identical to that of the immediately preceding “first” conjunct clause. (3.2) illustrates an empty expletive pronoun subject, while (3.3) illustrates an example of a third category. In this specific

instance, the empty pronominal subject occurs in a second conjunct clause and shares reference with the direct object of a preceding main clause, and does thus not count as a subject pronoun elided under coordination.

Now, these “other cases” have been stated by the YCOE corpus analysts to include tokens corresponding to what generative literature refers to as *pro*. Many of our cases thus represent genuine *pro*-drop. Once again, *pro* is defined as an empty element present whenever a finite verb has no overt subject, and this empty subject “is not coreferential with that of the previous parallel clause.”<sup>1</sup> In addition to cases of *pro*, this category includes cases where the empty pronoun may be co-referent with the subject of the immediately preceding “first” conjunct clause, but where some other factor interferes. Such complicating factors may include cases where the number changes, or where there is only partial co-referentiality, i.e. the empty pronoun shares reference with only part of the preceding subject. As will be shown in Chapter 4, instances are also included where the empty pronoun is co-referent with more than one single clause element, such as, for instance, a subject and an object, or a subject and an object of a preposition. These are instances which do not conform fully to the generative *pro*-category, but which will be treated here regardless.

The types of empty subjects exemplified in (3.1) and (3.2), however, will not be treated here. Empty coordinate subjects are not treated on basis of being idiomatic in OE, in the same way as they are still idiomatic in PDE. Empty subject-realizations of the expletive, or non-referential pronoun *hit* ‘it’ is also outside the scope of this work, which focuses on omission of referential subject pronouns. However, empty expletive subjects as a phenomenon is not unrelated to the subject matter of this study, as the existence in a language of non-referential pronouns has often been linked to the language’s status as a non-*pro*-drop language (Haugland 2007: 46). At this point, it could also be noted that the modern language does not sanction empty expletive subjects.

We are then left studying the “odd” cases of subject omission which often, but not necessarily always, correspond to the generative “small *pro*” category. It could also be reiterated here that the phenomenon under analysis goes counter to the grammatical rules of PDE, which generally requires overt subject pronouns, with a few exceptions (cf. also section

---

<sup>1</sup> Most of the information given here concerning YCOE and its use is chiefly paraphrased from the various reference manuals associated with the corpus, to be found at <http://www-users.york.ac.uk/~lang22/YCOE/YcoeHome.htm>. These include reference manuals for searching the corpus (<http://www-users.york.ac.uk/~lang22/YCOE/doc/corpussearch/CSRefToc.htm>) and for the syntactic annotations themselves (<http://www-users.york.ac.uk/~lang22/YCOE/doc/annotation/YcoeRefToc.htm>).

1.2). As also stated in Chapter 1, this is also true for the other Germanic languages, as well as French.

### 3.4 Corpus Linguistics

Data collection for this thesis has relied on corpus linguistics. Corpus linguistics as a method for linguistic research involves the electronic searching of data collections which have been specifically prepared for this purpose. De Smedt defines a corpus as “a collection of language materials (written or spoken) which are specially edited for use in language studies”.<sup>2</sup> The use of electronic corpora in linguistics dates back to the nineteen-forties, according to McEnery *et al.* (2006), yet it was only as personal computers gained widespread availability that “the exploitation of massive corpora became feasible” (McEnery *et al.* 2006: 3-4). This is of course due to the massive storage capacity of the modern computer. However, McEnery *et al.* point out that the corpus method had been utilised well before that time, by researchers using “shoeboxes filled with paper slips” as means of data storage. Regardless of how unrepresentative or skewed the corpus data may have been, McEnery *et al.* rightly state that the methodology of these researchers was “‘corpus-based’ in the sense that it was empirical and based on observed data” (McEnery *et al.* 2006: 3).

The most commendable feature of corpus linguistics as a methodology is that it allows the study of linguistic features without manually searching for tokens in texts or personally interviewing respondents. McEnery *et al.* list four main criteria for corpora, namely that the data should be authentic, machine-readable text “sampled to be representative of a particular language or language-variety” (McEnery *et al.* 2006: 5). They further define a corpus as “a body of naturally occurring language”, but also stress by referring to Leech (1992) that corpora are most regularly compiled with a specific purpose in mind (McEnery *et al.* 2006: 4). Such purposes could be to facilitate studies of the language of teenagers in London (the Corpus of London Teenage English) or to compare languages (the Oslo Multilingual Corpus). There are, of course, also many non-specialised or “standard” corpora, such as the British National Corpus.

In historical linguistics, it is a fact that researchers are naturally restricted to considering the data contained in the extant texts, as the case is in this study. There are no native speakers alive to act as informants, and as such, McEnery *et al.* (2006: 178) point out that studies of this kind have “always, in a sense, been corpus-based”. Indeed, they are no

---

<sup>2</sup> Collected from  
<http://gandalf.uib.no/lingkurs/webroot/index.php?page=corpus/collections&lang=en&course=lingmet>.

doubt also right in claiming that “[d]iachronic study is perhaps one of the few areas which can only be investigated using corpus data” (McEnery *et al.* 2006: 96). There should be little doubt, then, that extensive edited and searchable corpora offer a genuine advantage over manually searching the entirety of the existing physical corpus of the published Old English texts.

In the case of this study, access to and use of the YCOE, a syntactically tagged corpus containing the entire OE prose tradition has proven invaluable. As mentioned in Chapter 1, the type of subject omission investigated here is a fairly restricted phenomenon in OE, and as will be demonstrated in Chapter 4, the frequencies of clauses containing empty subjects are very low. Many of the texts under analysis are fairly long, and manually searching them for examples would likely take an excessive amount of time. Indeed, Susan Pintzuk has expressed her opinion that “it would be impossible to do much new on argument drop in Old English without using the YCOE” and that “it would take years to search the texts manually for examples” (personal communication May 2009). In order to illustrate this point, it may be mentioned that Sweet’s edition of *Cura Pastoralis* consists of 469 pages of prose containing approximately 3400 clause structures. Distributed quite evenly over these pages are no more than 37 instances of empty pronouns functioning as subject. It should be clear that manually searching for these 37 occurrences would be time-consuming, and such an approach would also entail considerable potential for inaccuracy and error.

As a consequence of this, the only feasible alternative to the corpus-method would have been to restrict the primary source material of this thesis to one – or even just part of one – single prose work. This would have been unfortunate in terms of the representativeness of the study, as it would have been very problematic to formulate generalisations of any kind based on such a small sample. It should also be considered that even investigating such a small sample would doubtlessly prove very time-consuming. Use of the YCOE has immensely simplified the process of data collection.

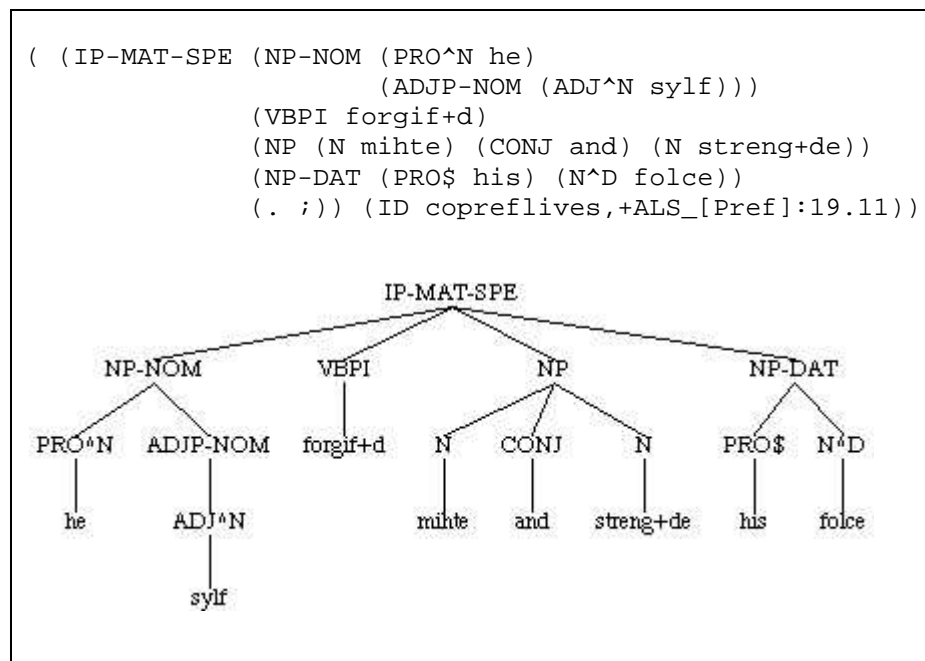
### **3.4.1 The York-Toronto-Helsinki Corpus of Old English Prose**

The YCOE is a syntactically annotated corpus consisting of 1.5 million words. It is the third in a series of historical English corpora, the other two being the Penn-Helsinki Parsed Corpus of Middle English II (PPCME) and the York-Helsinki Parsed Corpus of Old English Poetry. Contained within YCOE is every extant prose text from the Old English period. The text of the corpus is in the main identical to that of the University of Toronto’s Dictionary of Old English project, with only a few alterations for “foreign” characters such as *æ*, *þ* and *ð*.



According to the corpus analysts, errors have been corrected “silently”, and some minor changes have been made to word-division and capitalisation in order to facilitate syntactic and part-of-speech annotations. The goal of the corpus is to provide a method of easily searching for syntactic constructions in the OE texts. In terms of theoretical affiliation, the corpus can perhaps be said to somewhat partial to the framework of generative linguistics, particularly in terms of the syntactic labels used, and also in the format of the parsed syntactic tree structures with which the corpus operates. An illustration has been provided in figure 3.1, demonstrating the syntactic annotation of an OE main clause, represented both by the text as it appears in the corpus source and output files, and by a figure representing the same text in syntactic tree structure.

Figure 3.1: An example of the syntactic annotation in YCOE<sup>3</sup>



Provided below is the same text in “regular” format:

- (3.4) *he sylf forgifð mihte and strengðe his folce;*  
 he self gives power and strength to-his people-D  
 ‘He himself gives power and strength to his people.’  
 (ÆLS (Pref) 19)

<sup>3</sup> This illustration has been taken from the syntactic annotation reference manual associated with YCOE at <http://www-users.york.ac.uk/~lang22/YCOE/doc/annotation/YcoeRef.htm#examples>. This information was retrieved on 23 July 2010.

As can be seen in the table, each word in the corpus has been tagged syntactically, and arranged hierarchically in a syntactic tree structure where each single node is searchable by the program associated with the corpus. All labels present in the tree function as nodes. It may be observed that there are three main “levels” in the hierarchy, namely the clause-level, the phrase-level and the word-level.

The corpus does not feature an easily accessible search-interface like most modern corpora. Instead, the corpus is accessed by the CorpusSearch program. This program requires two separate types of input files which serve to instruct the corpus in what to search for, and also informing it of where it will find the requested information. The first is a command file, which normally would be a query file identifying the node or nodes to be searched and a specific query defining the relevant syntactic feature. The second is a source file in which the search is carried out. In our case, the source file would be a parsed corpus file containing a specific prose text. Finally, after the search has been finalised and tokens found, the CorpusSearch program creates an output-file detailing the identified citations.<sup>4</sup>

### 3.4.2 Data collection

In the case of this thesis, search parameters would have to be defined which would efficiently and reliably identify the relevant occurrences of empty subjects in the clause structure. Thus, a search file was written instructing the program to search in all *inflectional phrases* and to search for all instances of empty subjects which are not directly co-referent with the subject of the preceding clause, nor an example of an expletive subject. The search file is illustrated in (3.5) below.

(3.5) Search parameters for empty pronominal subjects  
node: IP\*  
query: \*pro\* exists

Our first concern is “setting the node”. This operation instructs the CorpusSearch program to search and print all instances corresponding with the query command, so long as their position in the underlying syntactic tree occurs below the identified node. As we are investigating sentential syntax, the node *IP\** has been set. The code *IP* indicates that the search should print all “hits” located in inflectional phrases. However, as a variety of types of inflectional phrases

---

<sup>4</sup> Again, the source of this paraphrased information can be found in the documentation of YCOE. Susan Pintzuk has also kindly instructed me in and recommended the search methods detailed in (3.6)–(3.9).

are identified in the tree structure, the wildcard command *\** was added, allowing the search to be carried out in all types of inflectional phrases. Other searchable nodes include, for instance *CP*, which facilitates searches to be made within complement phrases, while the *NP* node allows studies of the internal structure of noun phrases. The query command *\*pro\* exists* will ideally list all instances of empty pronouns tagged as *\*pro\**, which corresponds with the “other cases” discussed above. Thus, empty coordinate subjects, identified by the tag *\*con\**, and empty expletive subject pronouns, identified by the *\*exp\** tag, are excluded.

For purposes of comparison, it was also of interest to identify the frequency of clauses containing overt subject pronouns. A complementing search was carried out to identify such pronominal subjects. Similar to the search parameters shown in (3.5), this search also set the node to *IP\**. However, in order to achieve as precise results as possible, a more refined method of searching the corpus had to be used to search for these tokens. This was achieved by instructing the corpus to “tie” each distinct “part” of the query command to the one preceding it, with the first part of the query command initially being “tied” to the specified node. These “parts of the query” are referred to as *calls* by Susan Pintzuk. This method reduces the chances of the corpus providing hits in “bits of structure in different parts of the node” (personal communication March 2010). Additionally, the command *exists* was replaced by the commands *iDoms* and *iDomsOnly*. These terms are frequently used in connection with syntactic tree structures in generative linguistics, and basically mean that a specified node “immediately dominates” another node. The term “immediately dominates” indicates that there are no other nodes present separating the two in question in the syntactic tree structure. *iDomsOnly* would then indicate that a specified node immediately dominates only one particular node. In both cases, the dominating node is referred to as the *mother* while the “dominated” node is referred to as the *daughter*.<sup>5</sup> The following search was used to identify overt pronominal subjects in the corpus texts:

(3.6) Search parameters for overt pronominal subjects  
node: IP\*  
query: (IP\* iDoms NP-NOM\*) AND (NP-NOM\* iDomsOnly PRO^N)

Here, as is evident, the node is set to *IP\**, and is repeated in the first call, while the first call is reiterated in the second. This search file allows the CorpusSearch program to search within all

---

<sup>5</sup> These “familial” terms are related to the fact that the syntactic tree follows a family tree structure. Thus, terms such as *ancestors*, *descendants*, *mothers*, *daughters* and *sisters* are commonly seen.

inflectional phrases for instances where any nominative noun phrases are immediately dominated in the syntactic structure by any inflectional phrase, while any nominative pronouns are immediately dominated by a nominative noun phrase.

Another major benefit of this search method was that it could be modified to show the clausal distribution of the overt subjects. That is, the search would be able to show the details of the clause types in which the overt subject pronoun occurs. Combined with the occurrences of  $S_\emptyset$  analysed for this study, these figures for overt pronoun subjects allow for the creation of very detailed statistical overviews of the clausal distribution of  $S_\emptyset$  versus  $S_{\text{pron}}$  in OE. For this purpose, the following search was used:

(3.7) Search parameters for overt pronominal subjects in subordinate clauses

node: CP-THT\*/CP-ADV\*/CP-REL\*

query: (CP-THT\*/CP-ADV\*/CP-REL\* iDoms IP-SUB\*) AND (IP-SUB\* iDoms NP-NOM\*) AND (NP-NOM\* iDomsOnly PRO^N)

Here, a few changes have been made to the search. First, the CorpusSearch program has been instructed to search not within inflectional phrases, but within various types of complement phrases. *CP-THT* indicates that *hæt*-clauses should be searched, *CP-ADV* indicates adverbial clauses, while *CP-REL* indicates relative clauses. Again, the wildcard symbol instructs the program that the search is to be carried out in all variants of a clause type. These nodes had to be searched individually, as it would be impossible to procure results from these three nodes simultaneously. In terms of the query command string, it should also be observed that the calls have once again been “tied” to each other, as described above. The search command itself indicates that the node identifying the clause type in question should immediately dominate a node indicating a subordinate clause, and that subordinate clause should immediately dominate a nominative NP, while that nominative NP immediately dominates a nominative pronoun. With these search parameters, three searches for each corpus text produced exact figures for the distribution of  $S_{\text{pron}}$  in *hæt*-clauses, adverbial clauses and relative clauses.

In order to provide results for overt pronominal subjects in non-conjunct and second conjunct main clauses, the search parameters once more had to be modified slightly. The search used to find overt pronominal subjects in non-conjunct main clauses is illustrated in (3.8).

- (3.8) Search parameters for overt pronominal subjects in non-conjunct main clauses  
 node: IP-MAT\*  
 query: (IP-MAT\* iDomsNumber 1 !CONJ) AND (IP-MAT\* iDoms NP-NOM\*)  
 AND (NP-NOM\* iDomsOnly PRO^N)

As (3.8) shows, the node was set to *IP-MAT\** in this instance, instructing the program to search in all kinds of *matrix clauses*, a term corresponding to main clauses. As the case was in previous searches, the query command once again ties the first call to the node, but a new search command – *iDomsNumber 1* – is introduced. This command essentially means “immediately dominates as first daughter”, and is used in this case to identify cases where a main clause immediately dominates an element other than a conjunction, when that element is the first daughter of the “dominating” main clause mother node in the syntactic tree. The term “first daughter” would identify the first of several nodes immediately dominated by another node, and it is the tag *!CONJ* which indicates that this “first daughter” is not a conjunction. The exclamation point is used as a negative operator in the CorpusSearch program – meaning, basically, “not”. Additionally, as was also done in the preceding searches, the query string instructs the program to search within main clauses for nominative NPs which immediately dominate nominative pronouns in the tree structure.

The search parameters given in (3.9) were used to search for overt pronominal subjects in second conjunct main clauses.

- (3.9) Search parameters for overt pronominal subjects in second conjunct main clauses  
 node: IP-MAT\*  
 query: (IP-MAT\* iDomsNumber 1 CONJ) AND (IP-MAT\* iDoms NP-NOM\*)  
 AND (NP-NOM\* iDomsOnly PRO^N)

The only difference between (3.8) and (3.9) is that the negative operator *!* preceding *CONJ* has been removed, indicating that the search program should produce hits for instances where the first “daughter” of the matrix clause tag is a conjunction. That is, the first word in the clause should be a conjunction, and the main clause node should immediately dominate a nominative noun phrase, which again immediately dominates a nominative pronoun.

### 3.4.3 Problems in using the YCOE corpus

While YCOE has been an invaluable tool in the data collection process, certain issues and problems did present themselves. First, a minor problem was observed with the way YCOE

positions empty pronouns in the clause, which resulted in some highly unconventional word order patterns. As the corpus analysts themselves admit, the empty pronoun label has been inserted at the earliest possible position in the clause containing it. In very many cases, this position did not correspond with what the location of an overt pronoun in a corresponding structure would have been. Thus, the empty pronominal subjects had to be moved to what was considered the most likely position, according to my best judgment. This was done by analogy, and it is believed that in most of the cases, the position of  $S_{\emptyset}$  reflects quite accurately what the position of  $S_{\text{pron}}$  would have been, according to clause type. While the word order of OE admittedly was somewhat less restricted than that in PDE, it was certainly not entirely free, and features a clear V2 constraint. This means that when the corpus placed the \*pro\* label initially in a clause that would normally be verb-second, the corpus “analysis” was accepted. When the \*pro\* label was placed initially in a clause where it was felt that verb-initial verb order would be the better analysis, however, such as in an interrogative clause, or a clause featuring a hortative subjunctive structure, the empty pronoun was moved accordingly. Other than this brief example, though, no space will be dedicated here to discuss the various word order patterns in main and subordinate clauses – as this has been done excellently and exhaustively by, among others, Beck (2001) and Heggelund (2010). Suffice it to say that once the most likely location of the “subject gap” had been determined, this concern ceased to be an issue.

More serious, perhaps, is the fact that it came to my attention at quite a late stage in the process that the *exists* command used to search for empty subjects may be “bugged”, and also “may not restrict the search domain sufficiently” (Susan Pintzuk March 2010, personal communication). As such, there may be a risk that not all relevant examples have been caught by the search program. To ensure that this is not the case, alternative search parameters, recommended by Pintzuk, were used to run another series of searches of the corpus texts. These searches provided no hits which were not already part of my collected data, hopefully indicating that the data collection method used here is not too deeply flawed, and that the CorpusSearch program has not “missed” a substantial number of tokens. Still, it would be an oversight to fail mentioning that this risk does exist.

Incidentally, a less serious consequence of the fact that the *exists* command does not restrict the search sufficiently was evidenced when searches occasionally produced hits in citations containing the grapheme-sequence <pro>, resulting in hits containing no empty subjects, but which featured this sequence in proper nouns such as *Sempronius*, *Deprobane*, and also in similar cases. It must be said, though, that this is a minor problem which was

easily rectified when the time came for manual analysis of the collected data – at which time these instances were discarded from the analysed material.

### 3.5 Method of analysis

The citations found by searching the YCOE were entered into Filemaker Pro version 8.5, where they were analysed and classified according to a set of primarily syntactic, but also pragmatic, criteria. The syntactic criteria correspond to the structural variables – referenced in section 1.7 – which will form the basis for the quantitative analysis aimed at ascertaining the role of the syntactic environment in which the empty subjects occur. First, the citation itself was entered into a separate field. Secondly, a gloss version of the OE text was added to another separate field, followed by another field containing a more or less idiomatic PDE translation. These translations were, of course, not relevant to the results of the study, yet proved highly useful in the process of analysis, particularly in terms of correctly identifying the *antecedent* of the empty pronoun. Identifying the antecedent was not always a straightforward matter, and the addition of glosses and idiomatic translations served to simplify that process immensely. This topic will be treated in more detail in Chapter 4. In cases where there were doubts as to the identity of the antecedent, the immediately preceding context was also added to the database, also in a field of its own.

Now, the term *antecedent* is well-known in syntax, and is very commonly used in conjunction with pronouns. It is remembered that pronouns commonly recapitulate or “point back towards” nouns and other pronouns. The antecedent is the element which the pronoun recapitulates, i.e. the antecedent is the element to which a pronoun refers, whether nominal or pronominal. Whenever a pronoun has an antecedent, the pronoun agrees with its antecedent in number, person and gender. Not all pronouns have antecedents, but referential subject pronouns typically do.<sup>6</sup>

When a subject pronoun is empty, the antecedent takes on particular importance, because it is by identifying the antecedent that the meaning of the clause emerges. In our context, the antecedent has been defined as the last iteration in the text of the expression with which the empty subject shares reference. If the reference of  $S_\emptyset$  should happen to be cataphoric, it is the first occurrence of the co-referent expression which will be considered to be the antecedent.<sup>7</sup>

---

<sup>6</sup> In cases of situational reference, however, the reference is to something outside of the text itself. In such cases, it is very difficult to speak of an antecedent.

<sup>7</sup> Cf. section 4.7.2. for the terminology employed.

Subsequently, the citations were classified according to various syntactic criteria. First, attention was given to the instance of  $S_\emptyset$  itself. An entry was made in the database indicating the number of the empty pronoun, whether singular or plural. In many cases, distinguishing number was an easy matter, as the corresponding verbal inflections serve to disambiguate this quite efficiently. In most cases, a singular verb indicates that the empty pronoun is singular, and a verb inflected for the plural would correspondingly indicate a plural pronoun. This feature was, incidentally, also useful in determining the identity of the antecedent (cf. also section 4.9). Secondly, a field indicating the likely person of the empty pronoun was filled in, showing whether the instance of  $S_\emptyset$  represented a first, second or third person pronoun. Third, the clause type in which the empty subject occurred was identified. Here, a distinction was made between non-conjunct main clauses, second conjunct main clauses, relative clauses, adverbial clauses and *þæt*-clauses (cf. section 4.7.1).

Further, information identifying the grammatical function of the antecedent was entered into the database (cf. section 4.7.2). Here, a distinction was made between antecedents functioning as subject, object, object of a preposition, subject complement or a genitive in its clause. As some instances of  $S_\emptyset$  featured antecedents which did not correspond with these main categories, a “miscellaneous” category was established to cover the excess examples (see further discussion in section 4.7.2.5). An “indefinite” category was also established, on the grounds that not all instances of  $S_\emptyset$  feature a syntactic or overt antecedent, but rather rely on indefinite reference (see further discussion in section 4.7.2.6).

Finally, the clause type in which the antecedent occurs was identified for each citation (cf. section 4.7.3). Four main categories were established in this case also, distinguishing between antecedents occurring in preceding main clauses, following main clauses, preceding subordinate clauses and following subordinate clauses. Once again, categories representing “miscellaneous” and indefinite antecedents were created.

The database has also been provided with four distinct categories denoting the distance separating  $S_\emptyset$  and its antecedent, distinguishing distances of 1–3 words, 4–6 words, 6–10 words and 10+ words separating  $S_\emptyset$  and the antecedent.

### **3.5.1 Problems in using FileMaker Pro**

Overall, FileMaker Pro has served its function well. While there was a certain learning investment, the program is intuitive and accessible, and once the record layout had been composed, it was an easy matter to enter the appropriate information. Once all the citations had been entered and analysed, FileMaker Pro also facilitated easy searching, and simplified



the process of collecting statistical information in conjunction with, for instance, the compilation of tables. However, one central disadvantage to using FileMaker Pro is the fact that the program is not overly “intelligent”. Statistical problems occurred because it was assumed that FileMaker was capable of distinguishing categories and criteria “intelligently” to a greater extent than it actually was. When summing up frequencies for antecedent type, it was found that the total number of occurrences came to 107%. This happened because FileMaker could not distinguish between the tags “O” (object) and “OofP” (object of preposition) in a satisfactory manner. Instances having these tags were counted together, which means that the prepositional objects were actually counted twice. This was rectified by manually counting the occurrences, which admittedly is somewhat contrary to the purpose of compiling a database in the first place. However, this proved to be only a very minor obstacle. Attached is an example of a FileMaker record.

### **3.6 Problems of analysis**

#### **3.6.1 Are the selected texts representative?**

As detailed in Chapter 1, the selection of OE texts forming the basis for this study is fairly extensive, comprising 11 works – six representing the early OE period and five representing late OE. In terms of content and sheer text length, the text selection is in one way somewhat skewed in favour of the IOE period. Figures displayed in the YCOE output-files show that my corpus of eOE texts consist of a total 16 341 tokens, while the IOE texts consist of no less than 29 677 tokens.<sup>8</sup> However, despite the fact that the IOE subcorpus is nearly twice the length of the eOE part, it was believed beneficial to include works by as many distinct authors as possible in order to achieve as high a degree of representativeness as possible. As Heggelund (2010) points out, one should always consider the problematic fact that “[t]he dominance in Old English of certain authors, such as Alfred and Ælfric, means that individual style may influence the results of linguistic investigations” (Heggelund 2010: 44). As two of the IOE corpus texts are authored by Ælfric, it was judged important and necessary to expand the corpus as much as possible. Thus, it was determined that the discrepancy in length was offset by the fact that more distinct texts would increase the representativeness of the

---

<sup>8</sup> The corpus analysts generally define a token as a sequence containing one main verb, i.e. roughly the equivalent of a clause, although there are exceptions and variations to this rule.

selection. The fact that the IOE texts are considerably more voluminous is perhaps unfortunate, but is deemed here to be the lesser evil.<sup>9</sup>

As will also become apparent, though, it was observed that the eOE texts have much higher observed frequencies for  $S_{\emptyset}$  compared to the IOE texts, despite being nearly half the total length of the later texts. As will be seen in Chapter 4, the total number of occurrences of  $S_{\emptyset}$  in the eOE texts is 352, while the same number for the IOE texts is 99. This could lead to concluding that it was necessary to include more IOE texts in order to achieve enough tokens to allow for conclusions to be drawn. Also, it is believed that a greater number of texts causes a greater level of representativeness. Therefore, I do not consider it problematic that the IOE period has somewhat greater representation in terms of text length. In fact, it is thought that the selection of texts is quite thorough, especially when it is considered how much text can realistically be included as background for a work of this limited size and scope.

Also, statistical tests for significance will serve to compensate for the fact that the size of the samples is uneven. The statistical test used here is the chi-square test. Unless otherwise specified, the test in question will be a contingency table chi-square test (<http://faculty.vassar.edu/lowry/newcs.html?>). For instances where comparable observed frequencies were not available, the chi-square goodness of fit test has been used (<http://faculty.vassar.edu/lowry/csfit.html>). In cases where this test has been used, a note indicating this will appear in the text. Probability values have been considered statistically significant at the customary 5 percent level. Df=1 unless otherwise stated.

However, no matter the quality and well-roundedness of the selection, the problem arises whether it really is possible to make claims for OE based on this selection of texts alone – or indeed any selection of texts. Does the language contained in these texts in all cases represent “genuine”, representative OE? Any informants are certainly long dead, and the language they spoke seems decidedly alien to most speakers of the modern variant of the language. Additionally, as stated in the introduction, the extant OE texts represent the West-Saxon dialect, for the most part, meaning that the texts may be less representative of the language of Anglo-Saxon England as a whole. Also, one should consider the fact that the extant OE material is very limited in size, which can render it problematic to reach valid linguistic conclusions. Kohonen (1978) touches upon this, saying that “many crucial arguments may have to depend on only a few occurrences of a form” (Kohonen 1978: 75).

---

<sup>9</sup> However, it should also be recalled that a case can be made for considering GD as belonging to the eOE period. If so, the balance would be shifted by 5373 tokens, which would even out some of the differences.

Another problematic aspect of these texts is the fact that the majority of them are composed not in vernacular OE, but are translations or paraphrases of Latin originals. This problem was also mentioned in conjunction with the brief overview of empty subjects in Gothic (section 2.7.1), which, as recalled, features extensive use of empty subjects. In conjunction with this, it was mentioned that the syntax of the original Greek may have influenced the syntax evident in Wulfila's translation of the Bible. Given the possibility of such influence, it cannot be completely ruled out that the evident instances of  $S_\emptyset$  could be the result of a non-idiomatic translation from the Greek, which was a genuine *pro*-drop language. The exact same problem exists for OE due to its high volume of translations from Latin. Like Greek, Latin was a fully *pro*-drop language, and as such did not require overt subject pronouns. It is thus not impossible that the language of the original text may have had at least some influence on the authors and scribes, and thus also on the extant OE texts. It is difficult to exclude the possibility that instances of  $S_\emptyset$  present in translations from Latin originals may in fact reflect the Latin idiom, and not that of native OE. As the majority of the OE text tradition is based on translations from the Latin, this could constitute a problem in our context.

At this point, though, it should be noted that Kohonen (1978) states that "it is generally assumed that OE texts are remarkably independent of the syntax in the Latin originals" (Kohonen 1978: 74). Haugland (2007: 13) also states that the translations "do not generally strike the reader as slavish, gloss-like latinised renderings." She also points out that the translated prose generally bears great similarity to the prose seen in vernacular OE originals, such as the *Anglo-Saxon Chronicle*. However, she does point out that there are a few notable exceptions, "particularly in Bede [...] where we find deviant or awkward structures which may be directly dependent on the source text or indirectly due to the translation process" (Haugland 2007: 13). As shall be seen in Chapter 4, the syntax of Bede consistently shows higher percentages of  $S_\emptyset$  in most of the categories defined in the study. However, while "latinisms" of this type are likely present to some degree in the analysed texts, one would in most cases expect the scribes to correct usages that contrasted with what they felt was grammatically acceptable. Indeed, the notion that scribes would correct syntactic usages at odds with their own sense of grammaticality is strengthened by evidence from Gothic (2.7.1), where several instances have been observed of overt pronouns appearing where the Greek original would have featured a structure with an empty pronoun. As previously mentioned, Axel (2005) also reports that, in certain syntactic environments, OHG texts consistently insert  $S_{\text{pron}}$  in cases where Latin originals feature  $S_\emptyset$ . Overall, then, it will be

assumed here that the corpus of texts generally represents genuine and grammatical OE usage, although one should certainly remain critical, perhaps particularly in conjunction with a text such as *Bede*.

In the end, though, one should realise that, while the extant material may be problematic in aspects such as those mentioned above, it does contain all the information there is to be found about Old English, and must therefore by necessity be considered to represent the language in its entirety. In many ways, the problem regarding representativeness really becomes something of a moot point, given that these texts represent our only vantage point into the language of the Anglo-Saxons, and thus also the only possible way of drawing any kind of conclusions about the English language as it was being written and spoken toward the end of the first millennium CE.<sup>10</sup> That being said, though, it is certainly still useful to be conscious of the limitations and shortcomings of the preserved corpus of texts.

### 3.6.2 Could $S_{\emptyset}$ be a result of scribal error?

Another methodological problem is the possibility of scribal error. It cannot be excluded that some of the instances of  $S_{\emptyset}$  discussed in this study are actually representative of neither OE nor Latin, but rather of grammatical mistakes. This would mean that any such occurrences of  $S_{\emptyset}$  would be misleading if used as evidence in favour of a given argument concerning omitted subjects in OE, and they may therefore serve to skew the analysis somewhat. The possibility of such scribal errors is acknowledged by the YCOE corpus analysts, who state that the *\*pro\** category may encompass such “erroneous” tokens. As will also be seen in Chapter 4, Allen (2000) highlights the possibility of scribal error as a potential explanation for some of the instances of  $S_{\emptyset}$ . However, it is highly difficult to decide exactly which instances, if any, are caused by scribal error. When dealing with a phenomenon of such restricted distribution, it would be very difficult indeed to decide which instances are genuine and which are errors, and it seems nearly impossible to devise a system capable of rigorously distinguishing between the two. Given these difficulties, it is also nearly impossible to determine what the statistical consequences would be of including potential errors in our data material. Thus, it seems best to let the general stance taken in this work be that the analysed occurrences of  $S_{\emptyset}$  represents genuine tokens valid for research. This stance also reflects, for instance, the view held by Mitchell (1985), who, as recalled, refers to  $S_{\emptyset}$  as idiomatic in OE (cf. section 1.2).

---

<sup>10</sup> Heggelund (2010) points out that the written language is often more conservative than the spoken language, and that the extant texts thus may represent slightly older idiom than what was actually spoken at the time of composition. This is chiefly important for purposes of dating of texts, of course (Heggelund 2010: 43). The fact remains that the information regarding the language of the common man remains largely unknown.

However, we should consistently have in mind the possibility that some of the analysed instances may not, in fact, represent genuine OE usage – yet this appears to be one of the cases where it is impossible to reach a satisfactory final conclusion. In historical linguistics, it is sometimes necessary to acknowledge the fact that the extant material will have to suffice. Thus, in cases such as this, it is important to remain critical, but still allow analysis to be carried out, despite how conclusive evidence may not always be attainable.

### **3.7 Summary**

This chapter has been concerned with the method utilised in connection with the processes of data collection and analysis. The first part of the chapter introduced terminology and limited the scope of the study. Following that, corpus linguistics as a method for data collection was discussed, along with a presentation of the electronic corpus used to procure the data forming the basis for the analysis to follow in Chapter 4. In order to achieve as great a degree of transparency as possible, section 3.4.2 detailed the exact search parameters used to search for instances of empty and overt pronominal subjects in the York-Toronto-Helsinki Corpus of Old English Prose. Some problematic aspects associated with use of the YCOE were presented, before attention was given to the way the collected tokens were analysed. Finally, the last part of the chapter was dedicated to problems related to the analysis. It is hoped that this chapter provides insight into the processes of data collection and analysis, and will serve to make the following chapter, constituting the main part of the study, as coherent and accessible as possible.

## CHAPTER 4

### DATA AND ANALYSIS

#### 4.1 Introduction

The present chapter is concerned with the examples of empty subjects ( $S_\emptyset$ ) evident in the selected corpus of Old English texts. The aim of the chapter is twofold. First, it attempts to provide a brief overview of the relative distribution of  $S_\emptyset$ , and compare this with the distribution of overt pronominal subjects occurring in the same corpus texts. On this basis, an attempt is made to formulate statements on the historical development of empty subject pronouns in OE. Secondly, in-depth presentations, analyses and discussions of the characteristics of these empty subjects are given. Here, the instances of  $S_\emptyset$  will be discussed based on the various syntactic and pragmatic criteria hypothesised to be of relevance to the phenomenon (cf. sections 1.6, 3.5). Among these can be mentioned the types of clause in which  $S_\emptyset$  appears, the grammatical function of the antecedent co-referent with the empty subject, the clause-type in which the antecedent occurs and the textual “distance” between the empty pronoun and its antecedent. However, immediately preceding this, a section will be dedicated to two specific variants of  $S_\emptyset$  which are significantly more uniform than the majority of the examples discussed here. These are empty subject pronouns occurring in “imperative-like” hortative subjunctive main clauses and empty subject relatives.

#### 4.2 The relative distribution of $S_\emptyset$ in OE

As mentioned earlier, previous scholarship has established that OE subject pronouns were most commonly expressed, yet that the language displayed a somewhat greater tendency to omit such pronouns compared to the modern language. Preceding sections have also provided insight into the status of subject omission in other Old Germanic languages. It was seen that Gothic relied extensively on empty subject pronouns (section 2.7.1), and the same could also be said to a certain degree about Old High German (section 2.7.2). Several OHG texts utilised  $S_\emptyset$  to a notably high degree, particularly the *Benediktinerregeln*, which featured empty subject pronouns in 88.6% of the cases where the subject of a clause was pronominal. At this point, then, it is worth remembering Pogatscher’s claim that “missing subjects” represent a language feature common to all the early West-Germanic languages (Pogatscher 1901: 276-278). In light of this and the various other statements issued about  $S_\emptyset$  in OE (cf. sections 1.2–1.3, 1.6), it is very much of interest to quantify exactly how often this phenomenon occurs in OE texts.

Table 4.1 demonstrates the distribution of  $S_\emptyset$  in OE prose as represented by my corpus texts, and provides a basic comparison between empty and overt pronominal subjects. The table gives the number of occurrences of overt and empty pronominal subjects in the corpus texts and the relative frequency of  $S_\emptyset$ , expressed as a percentage of the total.

Table 4.1: Pronominal subjects in Old English prose:  $S_{\text{pron}}$  vs.  $S_\emptyset$

Texts	$S_{\text{pron}}$	$S_\emptyset$	
		n	%
Bede	3 948	133	3.3
Bo	3 897	26	0.7
ByrM	356	1	0.3
ChronA	363	7	1.9
CP	4 839	36	0.7
GD	5 296	15	0.3
LawAf	131	84	39.1
Or	2 303	65	2.7
WSCp	5 792	35	0.6
ÆCHom I	5 583	26	0.5
ÆCHom II	4 801	22	0.5
<b>Total</b>	<b>37 309</b>	<b>450</b>	<b>1.2</b>

The table shows that  $S_\emptyset$  must generally be said to be a very restricted phenomenon in OE. The overall frequency for  $S_\emptyset$  versus  $S_{\text{pron}}$  in the combined texts is 1.2%, a very low rate indeed. Excepting Alfred's *Laws*, which constitutes a rather special case, frequencies are generally low on a text-to-text basis, ranging from 0.3% (ByrM, GD) to 3.3% (Bede). The remarkably high percentage of 39.1%  $S_\emptyset$  in LawAf is mainly due to its high concentration of "imperative-like" hortative subjunctive structures, about which more will be said below. Let it be stated again however, that the omission of a subject pronoun in imperative structures is not an uncommon feature in OE, and does not represent the type of omitted subjects with which this work is primarily concerned. As will be seen below, these hortative subjunctives behave in a fashion very similar to imperatives, and the implications of the high concentration of such clauses in the *Laws* will be further discussed in section 4.5.

Aside from the *Laws*, then, the highest frequencies are seen in the OE versions of *Bede* and *Orosius*, at 3.3% and 2.7%, respectively. Behind *Orosius* comes the *Anglo-Saxon Chronicle*, which displays 1.9% empty subjects. Behind these texts, the percentages drop quite sharply down to the 0.7% of both *Boethius* and *Cura Pastoralis*. While the frequencies of 0.5% for both the first and second series of *Ælfric's Catholic Homilies* are certainly also low, it may perhaps be more interesting to note here the fact that the figure is identical for

both texts. The two volumes of the *Catholic Homilies* are quite extensive texts, and for them to show identical frequencies for  $S_{\emptyset}$  may suggest some level of systematicity – or otherwise a notable coincidence.

In this context, it may also be interesting to compare the figures for the OE version of *Boethius* with that of the OHG version. As table 4.1 demonstrates, the OE version features 0.7% empty pronouns in subject position, while the OHG version has 0.9%, as referenced in section 2.7.2. As the case also was with *ÆCHom I and II*, these frequencies are notably even, a fact which may perhaps serve to corroborate the notion that limited use of empty subjects was indeed a realisation of a language-feature common to early stages of the Germanic languages. It should be noted, however, that compared to the other texts studied in Eggenberger (1961) (cf. section 2.7.2) frequencies are consistently much lower for the OE texts studied here.

Indeed, one should always keep in mind that the frequencies overall are very low for this phenomenon in OE. As has been stressed repeatedly, pronominal subjects were fully lexicalised in the overwhelming majority of OE clause structures. However, the percentages indicate that the phenomenon is frequent enough to warrant closer examination. This is supported by the fact that every single text under analysis in this work contains examples of  $S_{\emptyset}$ .

### **4.3 The historical development of $S_{\emptyset}$**

It has been suggested at several stages in the course of this thesis that OE might represent a stage in a development in the English language toward losing empty subject pronouns. As recalled, overt subject pronouns are the norm only within the Germanic branch of Indo-European, while the majority of the world's languages allow empty pronouns to function as subject, regardless of whether those languages feature inflectional morphologies. Chapters 1 and 2 have also shown that overt subject pronouns constituted a competitive innovation in the Germanic languages, which ultimately supplanted the empty variant (sections 1.2, 2.7.2). Given this, it might be expected that there would be more examples of  $S_{\emptyset}$  present in the texts belonging to the early Old English period, as the late Old English period would represent a “further” step in a process leading toward the loss of empty subject pronouns. Table 4.2 compares the distribution of  $S_{\emptyset}$  in the early and late OE periods. The table gives the number of occurrences and the relative frequencies, expressed as a percentage of the total number of occurrences of pronominal subjects – overt and empty.



Table 4.2: S<sub>pron</sub> vs. S<sub>ø</sub> in eOE and IOE prose

Early OE texts	S <sub>pron</sub> n	S <sub>ø</sub> n	%	Late OE texts	S <sub>pron</sub> n	S <sub>ø</sub> n	%
Bede	3 948	133	3.3	ByrM	356	1	0.3
Bo	3 897	26	0.7	GD	5 296	15	0.3
ChronA	363	7	1.9	WSCp	5 792	35	0.6
CP	4 839	36	0.7	ÆCHom I	5 583	26	0.5
LawAf	131	84	39.1	ÆCHom II	4 801	22	0.5
Or	2 303	65	2.7				
<b>Total</b>	<b>15 481</b>	<b>351</b>	<b>2.2</b>		<b>21 828</b>	<b>99</b>	<b>0.5</b>

As the table demonstrates, the frequency for S<sub>ø</sub> is indeed higher in eOE, showing an overall frequency of 2.2% empty pronominal subjects, compared to the 0.5% observed in IOE. The difference between the two periods is statistically significant ( $\chi^2 = 241.87$ ,  $p \leq .0001$ ). However, there is considerable intertextual variation, and both the low overall and text-individual frequencies underline the fact that this is a very marginal phenomenon in both periods. Given these low frequencies, it will certainly be difficult to claim that great developments toward the loss of empty subjects have been made in the time separating the early and late OE periods, even though the difference between them is statistically significant. One will also notice that the absolutely highest frequencies for S<sub>ø</sub> are seen in the eOE period.

At this point, though, it should be stressed again that the statistics of LawAf are not strictly comparable to the others due this text's concentration of "imperative-like" instances of S<sub>ø</sub> mentioned above (also, cf. section 4.5). When this text is discounted, the overall frequency for eOE drops to 1.7%. However, the difference between the periods nonetheless remains statistically significant ( $\chi^2 = 148.27$ ,  $p \leq .0001$ ). It should be noted here, though, that these statistical tests only indicate that the differences in distribution of the two pronominal subject variants presented here are not random, but correlate with period. They do not indicate whether it is the time period or some other factor that is decisive in terms of explaining the differences. In principle, the differences between the two periods could simply be due to individual variation.

With LawAf excepted, then, the higher rates of S<sub>ø</sub> in eOE primarily stem from the relatively high frequencies in *Bede* and *Orosius*, at 3.3% and 2.7%, respectively. These frequencies are, in this context of generally low figures, notably higher than the highest frequency found in the IOE period, namely the 0.6% of the West-Saxon *Gospels*.

Now, except for the "high" frequencies of *Bede* and *Orosius* – and the misleadingly high result for Alfred's *Laws* – the frequencies are actually quite similar for both periods.

Byrhtferth's *Manual* and the two series of Ælfric's *Catholic Homilies* show frequencies of 0.3%–0.5%, which is very close to the frequency of 0.7% seen for both *Boethius* and *Cura Pastoralis*. The absolutely lowest frequencies in the entire corpus are the 0.3% of ByrM and GD, but in the case of ByrM, it should be noted that this figure in fact represents only a single instance of  $S_{\emptyset}$ , which makes it problematic to attach great value to a percentage value. Additionally, this single instance is a hortative subjunctive of the same variety which is seen so often in LawAf. It could also be remarked that the lowest frequencies seen in eOE texts, those of CP and Bo, are still higher than ÆCHom I and II, which besides ByrM, demonstrate the lowest percentages of the IOE period. Still, these differences are negligible.

#### **4.4 Introductory remarks on the analysis and categorisation of the instances of $S_{\emptyset}$**

While the research tradition on empty subjects in OE is certainly very tentative, the preceding sections on previous research have suggested that  $S_{\emptyset}$  in OE is unrestricted in terms of the clause-types in which it may appear. That is, empty subjects have been said to occur in all kinds of clauses, whether main or subordinate – as well as in the various subgroups of these categories. As we shall see, the data collected in conjunction with this work provides substantiation of these claims. Various other claims have also been made regarding the syntactic environments in which  $S_{\emptyset}$  occurs, which will be tested in the following sections. Thus, the ensuing sections will describe and analyse in some detail the instances of  $S_{\emptyset}$  collected on basis of the primary sources.

First, a demarcation will be made between those instances of  $S_{\emptyset}$  which fall into clearly defined groups and those which exhibit more variation. The first category, which is the most restricted in terms of observed tokens, will be addressed in sections 4.5 and 4.6. These sections present instances of  $S_{\emptyset}$  occurring in hortative subjunctive constructions and as the subject of relative clauses, respectively. As will be seen, the constructions discussed in sections 4.5 and 4.6 are uniform to the point where it may be asked whether they represent forms of expression that were idiomatic at some stage of the language. It might be an overstatement to refer to these constructions as idiomatic at the time of the extant OE texts, as the observed frequencies are very low, but they do occur with enough regularity in the data material to suggest that the constructions may be remnants of fixed expressions. The fact that they – relatively speaking – show a high degree of uniformity facilitates analysis and classification to a much greater extent than the remaining examples. Therefore, sections 4.5 and 4.6 will be considerably shorter than the sections dealing with the less uniform examples.

The examples which may defy easy classification according to construction type will be dealt with in the considerably longer sections 4.7–4.10. These subsections are by far the more extensive. The quantitative analysis of the role of structural variables proposed by scholars such as Pogatscher (1901) and Mitchell (1985) will take place in these sections, and will comprise both the bulk of this study. This is partly due to the fact that these instances are more numerous, at 267 observed instances compared to the combined 183 of the two categories mentioned above, but also because they demand more extensive treatment due being more heterogeneous.

#### 4.5 Hortative subjunctives

A considerable group of the instances of  $S_{\emptyset}$  under analysis occur in so-called hortative subjunctive structures. These instances of  $S_{\emptyset}$  combine with subjunctive verbs in structures functioning in a manner closely resembling imperatives. As recalled, imperatives constitute one of the few areas in which the modern language permits empty subject pronouns. Indeed, it is perhaps the only case where  $S_{\emptyset}$  is the favoured option in the modern language, as demonstrated by example (4.1) below. As (4.2) shows, imperative structures containing overt subject pronouns are also sanctioned in PDE, yet (4.1) would more than likely be the preferred realisation in most cases.

- (4.1) *Be quiet!*  
 (4.2) *You be quiet!*

The situation in OE resembles this, in that variation is seen between the two options, yet the situation in OE is somewhat more complex, in that use of subject pronouns varies between positive and negative imperatives. The negative variant of imperative construction, introduced by the negative adverb *ne* ‘not’, normally prefers overt pronominal subjects (Mitchell 1985: 374–375, 383–384), as seen in (4.3).

- (4.3) *Ne wep þu:*  
 not weep-IMP.SG you  
 ‘Do not cry.’  
 (ÆCHom I 425.187)

Positive imperatives, on the other hand, normally feature empty pronominal subjects, similarly to the situation in PDE. Still, overt subject pronouns are also observed to occur,

albeit less often (Mitchell 1985: 374–375). Below are given two examples of OE positive imperatives, where (4.4) contains an empty pronominal subject and (4.5) the overt variant.

- (4.4) *Gif ðu hælend crist sy. Gehæl [S<sub>ø</sub>] ðe and us;*  
if you savior Christ be, save-IMP.SG [you-SG] you-RFLX and us  
'If you are the Saviour Christ, save yourself and us.'  
(ÆCHom II 146.249)
- (4.5) *Ondswarede he him: Gif he Godes man sy, fylgað ge him.*  
answered he them: if he God's man be, follow-IMP.PL you-PL him  
'He answered them: If he is God's man, follow him.'  
(Bede 100.23)

In terms of syntactic features, it is observed that positive imperative clauses are verb-initial, with the subject in second position when the pronoun is overt. By analogy, it can be assumed that this is also the likeliest location of the empty subject, reflected by its placement in example (4.4). According to Mitchell (1985: 375), however, overt subject pronouns may in some cases also occur in initial position. The negative variant is slightly different. As said above, these structures are introduced by the negative adverb *ne*, meaning that the verb is pushed to the second position and the subject to the third. However, if *ne* is considered to be a so-called *clitic* (cf. for example Koopman 1997 and van Kemenade 1987), and therefore not part of the clause structure, the word order remains identical to that of positive imperatives.

Morphologically, the imperative mood in OE has distinct inflectional endings only in the second person singular (ZERO) and plural (*-ap*). Quirk and Wrenn (1976) thus maintain that the “imperative proper only exists in the second person singular and plural,” yet allow that there may exist a very rare first person plural form (Quirk and Wrenn 1976: 85).

For exhortations in the third person, the subjunctive mood takes the function of the imperative. It should be commented, though, that while the subjunctive mood is mainly employed to express commands and exhortations in the third person, structures are also observed where the subjunctive has been used in the second person, as in example (4.6) below.<sup>1</sup> The result is constructions which resemble imperatives to a significant degree. These structures should, however, not be identified with genuine imperatives, which Mitchell stresses by “unrepentantly” referring to this verbal category as the *jussive subjunctive* (Mitchell 1985: 373).

---

<sup>1</sup> It is difficult to determine why the subjunctive mood is used for some exhortations in the second person instead of the imperative proper. It could, however, be speculated that subjunctive structures were considered a weaker form of exhortation, thus providing a veneer of politeness for what is indisputably a command.

Hortative subjunctives resemble imperatives in several ways. Semantically, they express commands, exhortations and encouragements. Syntactically, they often have a clause-initial verb, while the subject – and presumably also any possible empty subject pronoun – is postverbal, and typically located to the immediate right of the verb.<sup>2</sup> The empty subjects occurring in hortative subjunctive structures can thus arguably be said to represent a syntactic “behaviour” which is reminiscent of that of empty subjects occurring in genuine imperative structures. For an illustration of the syntactic similarities between imperatives and hortative subjunctives, examples (4.6) and (4.7) have been provided.

(4.6) *Drihten. gehæle [S<sub>ø</sub>] me*  
 lord, save-SBJCT [you] me  
 ‘Lord, save me.’  
 (ÆCHom I 323.163)

(4.7) *gif his hwa sie lustfull mare to witanne, sece [S<sub>ø</sub>] him þonne self þæt.*  
 if of-this anyone be desirous more to learn, seek-SBJCT [he] him then self that  
 ‘if anyone desires to learn more of this, let him seek it himself.’  
 (Or 56.11)

Example (4.6) exemplifies a very similar structure to that of (4.4). Aside from the introductory adverbial preceding the imperative clause in (4.4) and the vocative in (4.6), the only discernible difference between the two is the imperative ZERO-ending of the verb in (4.4) and the subjunctive *-e* ending of the verb in (4.6). In both cases the main clause features verb-initial word order, with S<sub>ø</sub> occurring in the second position. The semantic similarities should be apparent as well, as both are an imposition to provide salvation. Example (4.5) likewise bears great similarity to (4.7), as both clauses are introduced by an adverbial of condition, followed by an imperative or hortative main clause. Again the difference between the two is seen in the contrast between the imperative plural *-aþ* ending of (4.5) and the subjunctive *-e* ending of (4.7).

The structure seen in (4.4), (4.5) and (4.7), where an introductory adverbial clause of condition is followed by an imperative or subjunctive main clause, is commonly observed for both variants. Thus, except for inflectional differences, nothing seems to indicate that empty

---

<sup>2</sup> A possible exception to this rule is seen in constructions with the indefinite *mon* ‘one’, which is sometimes seen to occur in the final position, exemplified in *XXX scillinga geselle him mon* ‘one should give him 30 shillings’ (LawAf 44). Several examples of this variety are seen in LawAf. However, it should be mentioned here that *mon* exhibits characteristics of both pronouns and full NPs. For instance, *mon* can appear in clause-positions where pronouns normally do not occur. This has caused some to consider *mon* a nominal element. The stance taken here will be that of van Bergen (2000), however, who concludes that *mon* must be considered a pronominal subject (van Bergen 2000: 116).

subjects in subjunctive clauses should be treated as fundamentally different to those in imperative clauses.

In the examples of hortative subjunctives above, the empty pronoun has been analysed as occurring in the second, postverbal position. Similarly to imperatives featuring empty pronominal subjects, the decision to analyse  $S_{\emptyset}$  as occurring to the right of the verb in hortative subjunctives is made by analogy, as overt subject pronouns in this type of construction also appear postverbally, as illustrated in example (4.8). Thus, the imperatives and the hortative subjunctives can be said to be syntactically comparable in that they feature verb-initial verb order, and have the subject position – unfilled or otherwise – in the position to the immediate right. It may be observed that no instances of negative hortative subjunctives featuring  $S_{\emptyset}$  have been found in the texts under analysis in this work.

- (4.8) *Monnes cinban, gif hit bið toclofen, geselle mon XII scillinga to bote.*  
 man's cheek-bone, if it be split, give-SBJCT one twelve shillings to compensate  
 'If the cheekbone of a man is split, one should give twelve shillings in compensation.'  
 (LawAf 50.1)

Table 4.3 below depicts the distribution of hortative subjunctives featuring  $S_{\emptyset}$  in the various texts analysed here, as compared to the total number of instances of  $S_{\emptyset}$  present.

Table 4.3: Hortative subjunctives compared to total occurrences of  $S_{\emptyset}$ , according to texts

Text	Total $S_{\emptyset}$	$S_{\emptyset}$ -subjunctives	
	n	n	%
Bede	133	2	1.5
Bo	26	2	7.7
ByrM	1	1	100.0
ChronA	7	0	0.0
CP	36	7	19.4
GD	15	4	26.7
LawAf	84	84	100.0
Or	65	5	7.7
WSCp	35	16	45.7
ÆCHom I	26	4	15.4
ÆCHom II	22	2	9.1
<b>Total</b>	<b>450</b>	<b>127</b>	<b>28.2</b>

As the table shows, an entire 127 of the 450 analysed citations containing  $S_{\emptyset}$  occur in hortative subjunctive structures. This corresponds to a percentage of 28.2% of the total number of tokens. There is, as seen, considerable inter-textual variation. In some texts, such

as *Boethius* and *Orosius*, only 7.7% of the occurrences of  $S_{\emptyset}$  are found in hortative subjunctives, while in other texts, the figure is higher. Generally, however, the frequencies are well below 30%. The great exceptions are the *West-Saxon Gospels*, at 45.7% and of course Alfred's *Laws*, where all the registered occurrences of  $S_{\emptyset}$  are in fact hortative subjunctives. This is also true of the single token in Byrhtferth's *Manual*. It is recalled that LawAf was remarkable in that an entire 39.1% of its pronominal subjects were empty, a figure no other text comes near to equalling. The high figure for both  $S_{\emptyset}$  and  $S_{\emptyset}$ -subjunctives in LawAf is explainable by the format and content of the text. As a collection of laws, this text takes the form of a series of admonishments, where the subjunctive clause sets out the required punishment when certain conditions have been met, as seen in (4.9–4.11).

- (4.9) *Gif him mon aslea oþer eare of, geselle [ $S_{\emptyset}$ ] XXX scillinga to bote.*  
 if him one strikes either ear off, pay-SBJCT [he] 30 shillings to compensate  
 'If anyone cuts an ear off, he must pay 30 shillings to compensate.'  
 (LawAf 46.159)
- (4.10) *Gif se hlyst oðstande, þæt he ne mæge gehieran, geselle [ $S_{\emptyset}$ ] LX scillinga to bote.*  
 if the hearing is-stopped, that he not may hear, pay-SBJCT [he] 60 shillings to compensate  
 'If the hearing is damaged, so that he may not hear, he must pay 60 shillings in compensation.'  
 (LawAf 46.1.160)
- (4.11) *Gif mon oðrum þæt neb ofaslea, gebete [ $S_{\emptyset}$ ] him mid LX scillingum.*  
 if one another the nose off-strikes, compensate-SBJCT [he] him with 60 shillings  
 'If anyone cuts the nose off another man, he must compensate him with 60 shillings.'  
 (LawAf 48.163)

Extensive parts of LawAf take the form of consecutive occurrences of such sentences. There can be little doubt that the reliance on this type of subjunctives has contributed greatly to the remarkably high figures for  $S_{\emptyset}$  in this text – in fact, there would have been no such occurrences without these third person exhortations. It may even be said that the genre of legal documentation, with its propensity for “commands”, facilitates the phenomenon of “subjectless” sentences to a much greater degree than texts belonging to other genres. Incidentally, this reasoning also explains the high frequency of  $S_{\emptyset}$ -subjunctives in the *West-Saxon Gospels*, as most of these represent instances where Christ instructs his disciples.

As seen above, the high occurrences of  $S_{\emptyset}$ -subjunctives in some texts create statistics which are not directly comparable with the other texts forming the basis of this study. For this

reason, once discussed here, hortative subjunctives will be disregarded in the ensuing sections.

While it is certainly interesting to observe how many of the total number of  $S_{\emptyset}$  occur in hortative subjunctives, it would undoubtedly be more useful to provide an overview of the ratio between empty and overt subjects in such structures. Such an investigation lies outside the scope of this work, yet a basic overview has been provided, showing the distribution of empty and overt pronominal subjects with three of the most frequently occurring hortative subjunctive verbs in LawAf. This overview is presented in table 4.4 below. The statistics presented in the table were procured by searching the Dictionary of Old English Web Corpus for occurrences in LawAf of the subjunctive forms *bete(n)*, *gielde/gelde* and *selle*. Matches were then sorted according to whether they featured nominal or pronominal subjects, before a distinction was made between overt and empty pronominal subjects. The indefinite pronoun *man* ‘one’ has been counted as a pronominal subject in this context.

Table 4.4:  $S_{\text{pron}}$  vs.  $S_{\emptyset}$  with some hortative subjunctive verbs in LawAf

Verb	$S_{\text{pron}}$		$S_{\emptyset}$	
	n	%	n	%
<i>Betan</i> ‘compensate’	0	0.0	38	100.0
<i>Geldan/gieldan</i> ‘give’	2	15.4	11	84.6
<i>Sellan</i> ‘give’	12	40.0	18	60.0
<b>Total</b>	<b>14</b>	<b>17.3</b>	<b>67</b>	<b>82.7</b>

As the table shows, the frequencies for  $S_{\emptyset}$  are very high with these subjunctive verbs. The overall frequency of 82.7% is notably high, particularly given the low general frequencies for  $S_{\emptyset}$ , as seen above. The highest frequency for  $S_{\emptyset}$  is seen in combination with the verb *betan* ‘compensate’, a verb which appears 38 times in the hortative subjunctive and which is empty in all of the cases. The lowest, at 60%, is seen with *sellan* ‘give’. These high frequencies, along with the high levels of uniformity observed in these instances of  $S_{\emptyset}$ , beg the question of whether this variant of  $S_{\emptyset}$  perhaps should be considered to be a fixed phrase, or a remnant of such a phrase. It should be remembered that while LawAf represents the eOE period, these are in many cases ancestral Germanic laws which have likely been conveyed orally for a very long time prior to the conception of the text version. As such, it could be speculated that these laws have preserved evidence of an archaic form of expression which to a great extent preferred the omission of subject pronouns occurring with hortative subjunctive verbs. Of the



14 instances of  $S_{\text{pron}}$  here, the indefinite *mon* ‘one’ actually represents 13 instances. The only exception, illustrated below, is an instance with *he* functioning as subject.

- (4.12) *Gif he losige, & hine mon eft gefo, forgielde he hine self a be his weregilde*  
 if he escape-SBJCT and him one after capture, pay-SBJCT he him self according to his  
 wergeld  
 ‘If he escapes, and one captures him afterwards, let him pay according to his wergeld.’  
 (LawAf 7.1)

This means, of course, that if *mon* ‘one’ is analysed as a full NP, as many are likely to do (cf. footnote 2 above),  $S_{\emptyset}$  is virtually categorical in these cases.

#### 4.6 *Hatan* + empty subject relative pronouns

Another distinct subgroup of the empty subjects under analysis here is that of empty relative pronouns functioning as subject in a relative clause in conjunction with the verb *hatan* ‘be called’. Mitchell refers to this practice as one of several variants of “naming constructions” (Mitchell 1985: 533, 616-619). Two such instances are exemplified below:

- (4.13) *He gesette under him gingran casere, [S<sub>∅</sub>] Maximus wæs haten,*  
 he placed under him younger emperor, [who] Maximus was called  
 ‘He placed under him a younger emperor, called Maximus.’  
 (Or 146.20)
- (4.14) *þær hine afedde sum eawfæst munuc [S<sub>∅</sub>] Romanus hatte. þreo gear.*  
 there him fed some pious monk [who] Romanus called-was, three years  
 ‘A pious monk called Romanus fed him there for three years.’  
 (ÆCHom II 92.23)

This usage is clearly at odds with that of PDE, as at least the “standard” varieties of the modern language require subject relative pronouns to be fully overt. Object relatives, on the other hand, may be quite freely omitted in PDE. One should also note the form *hatte*, an archaic variant of this naming construction (Mitchell 1985: 616-617). For some texts, a considerable number of the instances of  $S_{\emptyset}$  found in the analysed texts belong to this category, while it is completely absent in others, as table 4.5 below shows. The table demonstrates the distribution of empty relative subjects ( $S_{\emptyset, \text{rel}}$ ) as they appear in the texts under analysis, compared to the total number of instances of  $S_{\emptyset}$ .

Table 4.5: The distribution of  $S_{\emptyset,rel}$  according to texts

Text	Total $S_{\emptyset}$		$S_{\emptyset,rel}$	
	n	%	n	%
Bede	133		23	17.3
Bo	26		2	7.7
ByrM	1		0	0.0
ChronA	7		0	0.0
CP	36		0	0.0
GD	15		0	0.0
LawAf	84		0	0.0
Or	65		28	43.1
WSCp	35		1	2.9
ÆCHom I	26		1	3.8
ÆCHom II	22		1	4.5
<b>Total</b>	<b>450</b>		<b>56</b>	<b>12.4</b>

As is evident from the table, a number of the tokens for  $S_{\emptyset}$  in the corpus texts represent instances of  $S_{\emptyset,rel}$ . Of the total instances of  $S_{\emptyset}$  in *Orosius*, for example, 43.1% are empty subject relatives. This frequency is notably higher than the frequencies observed in the other texts. The frequency of 17.3% in *Bede* is also quite high. Five of the texts feature no instances of  $S_{\emptyset,rel}$  whatsoever, though. As a result of this, the overall relative frequency for  $S_{\emptyset,rel}$  in the texts under analysis is quite low, at 12.2%. This frequency is notably lower than that observed for  $S_{\emptyset}$  occurring in hortative subjunctives above. As the table demonstrates, the collected tokens of  $S_{\emptyset}$  represent other uses of empty subjects than empty subject relatives in most cases.

It should be mentioned, however, that the high frequencies for  $S_{\emptyset,rel}$  in *Orosius* and *Bede* may not be exclusively motivated by syntax. In fact, these high frequencies may perhaps be better explained by textual factors, the most relevant of which being that *Orosius* and *Bede* are historical narratives, in which a great deal of space is dedicated to introducing the names of characters and places. This leads to higher concentrations of “naming” constructions than in other types of text, which in turn leads to higher concentrations of  $S_{\emptyset,rel}$ . On this background, though, it is notable that ChronA, which also is a historical narrative, features no examples of  $S_{\emptyset,rel}$ .

It should be mentioned here that while table 4.5 is certainly useful in determining how many of the observed instances of  $S_{\emptyset}$  are empty relatives occurring in combination with a form of *haten* or *hatte*, this table does *not*, however, depict the relative frequencies of  $S_{\emptyset,rel}$  versus what might be called  $S_{pron,rel}$ , i.e. overt relative pronouns functioning as the subject of a relative clause. Although such a comparison would no doubt be much more interesting, it has

proven beyond the capacity of this work to compile an overview of the situation regarding empty and overt relative subject pronouns. Still, a good way of supplying at least an indication of the ratio between the two would be to present a short study aimed at identifying the frequency of  $S_{\emptyset,rel}$  versus  $S_{pron,rel}$  in conjunction with the verbs *hatan* or its archaic variant *hatte*, as they occur in the analysed texts. Although it has proven too time-consuming to conduct such an investigation here, an example of such a structure is given in (4.15) below.<sup>3</sup>

- (4.15) *Wæs he sended from Westseaxna cyninge, se wæs haten Cwichelm*  
 Was he sent by West-Saxon king-G, who was called Cwichelm  
 ‘He was sent from the king of the West-Saxons, who was called Cwichelm’  
 (Bede 122.9)

Additionally, (4.16) and (4.17) below constitute parallel examples demonstrating contrasting usages of  $S_{\emptyset,rel}$  and  $S_{pron,rel}$ , respectively.

- (4.16) *& on his stowe gehalgode Seaxwulf biscop, se wæs timbrend & abbud þæs mynstres,*  
 and in his place hallowed Seaxwulf bishop, who was builder and abbot of-the monastery  
*[S<sub>∅</sub>] æt Medeshamstede is cweden, in Gyrwa londe.*  
 which at Medeshamstede is called, in Gyrwa’s land.  
 ‘and in his place hallowed Seaxwulf bishop, who was the founder and abbot of the  
 monastery, which is called Peterborough, in the Gyrwa’s land.’  
 (Bede 280.24)

- (4.17) *Wæs ðis gefeoht geworden on þære mæran stowe þe cweden is Degsastan.*  
 was this battle happened on the famous place which called is Dawston  
 ‘This battle took place in the famous place which is called Dawston.’  
 (Bede 92.17)

#### 4.7 Other uses of $S_{\emptyset}$

As previously mentioned, the data in the ensuing section is less homogeneous than the relatively clearly demarcated tokens discussed in sections 4.5 and 4.6. Even so, numerous patterns emerge even here – and some quite strikingly so. These patterns are based upon the various linguistic variables used in analysing the data collected, and while the end result may not lead to the formulation of a rule for subject omission in OE, it might at the least provide an indication of the environments in which such omission occurred. As mentioned above, the following sections will provide quantitative testing of the variables proposed in the existing literature, and each variable will form the focus of several sub-groups in the following text.

<sup>3</sup> Mitchell (1985ii: 88), however, points out that we cannot be sure whether *se* functions as a relative or a demonstrative pronoun in examples such as (4.15). This problem will not be addressed here.

These variables include many distinct aspects of  $S_\emptyset$ , such as the various clause types in which the phenomenon occurs, the type of antecedent with which the empty pronoun shares reference, the textual “distance” between antecedent and  $S_\emptyset$ , and so forth.

#### 4.7.1 The clausal distribution of $S_\emptyset$

A main issue in the literature on subject omission in early Germanic, and also in that directly pertaining to OE, has been the clausal distribution of the empty pronoun. As indicated at several points in the course of this work, the existing studies of this phenomenon, while certainly few and somewhat tentative, have suggested that the distribution of  $S_\emptyset$  must have been free, in that it could occur in all clause types. However, as observed in sections 1.3 and 2.8, only slight space is dedicated to the phenomenon, and very little quantification is done. As was also stated in these sections, it has been a central focus of this thesis to remedy this by quantifying the phenomenon to as large a degree as possible.

Thus, this section seeks to provide a picture of the relative distribution of  $S_\emptyset$  in the various clause types in OE. The following sections will present and discuss the tokens of  $S_\emptyset$  as they appear in the non-conjunct main clauses, second conjunct clauses, relative clauses, *þæt*-clauses and adverbial clauses of the analysed texts. For purposes of comparison, the frequencies for  $S_\emptyset$  will be contrasted with statistics for the distribution of  $S_{\text{pron}}$  in the same clause types throughout.

Finally, it should perhaps be reiterated that instances of hortative subjunctive structures featuring empty subjects and relative clauses containing empty subject relatives have been omitted from this part of the study. This means that the instances of  $S_\emptyset$  in relative clauses given in the table below do not represent cases of omitted subject relatives, as discussed in section 4.6, but rather empty subjects in clauses introduced by an object relative pronoun.

To introduce this component of the study, table 4.6 has been provided. The table is well-suited for introductory purposes, as it contains a concise summary of the data presented in the ensuing sub-sections (4.7.1.1–4.7.1.5), contrasting the total occurrences of  $S_\emptyset$  and  $S_{\text{pron}}$  in the corpus texts, according to clause type. While a more detailed account will be offered in the following sections, where the data for each clause type will be presented also according to the distribution in the different corpus texts, it is believed that a few initial comments may be made on basis of table 4.6 alone.

Table 4.6: The distribution of  $S_{\text{pron}}$  vs.  $S_{\emptyset}$  according to clause type

Clause type	$S_{\text{pron}}$		$S_{\emptyset}$	
	n	%	n	%
Main clauses	10 546	99.5	52	0.5
2 <sup>nd</sup> conjuncts	5 637	97.6	139	2.4
Relative clauses	2 779	99.9	4	0.1
Adverbial clauses	7 221	99.5	39	0.5
<i>þæt</i> -clauses	4 414	99.3	33	0.7
<b>Total</b>	<b>30 597</b>	<b>99.1</b>	<b>267</b>	<b>0.9</b>

As the table shows, the overall distribution of  $S_{\emptyset}$  versus  $S_{\text{pron}}$  is very low indeed, at a frequency of 0.9%. Granted, some variation is evident, with  $S_{\emptyset}$  being much more frequent in second conjunct clauses, yet this low overall frequency is somewhat surprising given the claims some scholars have made about empty subjects in OE. Indeed, the overall frequency of 0.9% really should be described as no more than negligible, and while it has been reiterated at numerous occasions that “subjectless” clauses in OE should be regarded as the exception rather than the norm, this is a notably low figure. This is especially true given the claims made by, for instance, Baker (2007), Mitchell (1985), Traugott (1992) and van Gelderen (2000) as referenced in section 1.2 – who have stated that non-expression of the subject should be regarded more or less as a regular feature of the language. Based on the figures of table 4.6, it must be concluded that these scholars may have made their claims on uncertain grounds, especially when it is considered that empty subjects in OE have not been systematically studied before. These conclusions remain valid also if one considers the frequencies for  $S_{\emptyset}$  in the specific clause types. For instance, the overall figure of 0.5%  $S_{\emptyset}$  in main clauses certainly does not support the concept that “pro-drop is quite common” in OE (van Gelderen 2000: 149). The same is true also for relative clauses (0.1%), adverbial clauses (0.5%) and *þæt*-clauses (0.7%). The only clause type which can be said to provide some evidence in favour of the claim that  $S_{\emptyset}$  is a regular part of the language is second conjunct clauses, although it must be said that even here, the frequency is very low.

#### 4.7.1.1 $S_{\emptyset}$ in main clauses

Attention will now be turned to a more detailed overview of the distribution of  $S_{\emptyset}$  and  $S_{\text{pron}}$  in non-conjunct main clauses according to the texts in which they occur. First, three such main clauses are exemplified – and as the examples illustrate, they may be preceded by a subordinate clause (4.18), (4.19), yet may also stand alone as a full sentence (4.20).

- (4.18) *Ɔa he Ɔa Wigheard to Rome becwom, ær Ɔon he to biscophade becuman meahthe,*  
 when he then Wigheard to Rome came, before that he to bishophood become might  
*[S<sub>ø</sub>] wæs mid deaðe forgripen,*  
 [he] was with death afflicted  
 ‘When he, Wigheard, came to Rome, he died before he could ascend to bishophood.’  
 (Bede 248.14)
- (4.19) *Gif he nære swutelice hreoflig.*  
 if he not-was manifestly leprous,  
*[S<sub>ø</sub>] wære Ɔonne be his dome clæne geteald.*  
 [he] was-SBJCT then by his judgment clean considered  
 ‘If he was not manifestly leprous, he was then considered clean, by his judgment.’  
 (Or 243.62)
- (4.20) *[S<sub>ø</sub>] Wearð Ɔa fordrifen on an iglond ut on ðære Wendelsæ.*  
 [he] became then away-driven to an island out in the Mediterranean-sea  
 ‘He was then driven to an island in the Mediterranean.’  
 (Bo 115.22)

While the data presented in table 4.6 reveals the most crucial information pertaining to  $S_{\emptyset}$  in main clauses, it is no doubt also of interest to provide a more detailed account of the distribution of  $S_{\emptyset}$  in main clauses according to the corpus texts, given the inter-textual variation demonstrated for the overall data (cf. table 4.1). This information is provided in table 4.7, below.

Table 4.7:  $S_{\text{pron}}$  vs.  $S_{\emptyset}$  in main clauses according texts

Texts	$S_{\text{pron}}$		$S_{\emptyset}$		Total	
	n	%	n	%	n	%
Bede	995	96.6	34	3.3	1 029	100.0
Bo	1 143	99.6	5	0.4	1 148	100.0
ByrM	138	100.0	0	0.0	138	100.0
ChronA	43	100.0	0	0.0	43	100.0
CP	943	99.6	4	0.4	947	100.0
GD	1 157	100.0	0	0.0	1 157	100.0
LawAf	25	100.0	0	0.0	25	100.0
Or	455	99.8	1	0.2	456	100.0
WSCp	2 173	99.9	3	0.1	2 176	100.0
ÆCHom I	1 857	99.9	2	0.1	1 859	100.0
ÆCHom II	1 617	99.8	3	0.2	1 620	100.0
<b>Overall</b>	<b>10 546</b>	<b>99.5</b>	<b>52</b>	<b>0.5</b>	<b>10 598</b>	<b>100.0</b>

As the table makes clear, the distribution of  $S_{\emptyset}$  in main clauses is extremely low in nearly all the texts. In fact, only seven of the eleven corpus texts feature  $S_{\emptyset}$  in main clauses, and for

most of those that do, the frequencies range between only 0.1% and 0.4% of the total. The sole exception to this is *Bede*, which features an empty pronominal subject in 3.3% of its non-conjunct main clauses containing pronominal subjects. The difference between *Bede* and the other texts combined is statistically significant at the five percent level ( $\chi^2=178.44$ ,  $p\leq.0001$ ). Also, if the instances of  $S_\emptyset$  are considered without reference to the instances of  $S_{\text{pron}}$ , it becomes apparent that the 34 instances in *Bede* comprise no less than 65.4% of the total number of occurrences of  $S_\emptyset$  in this clause type. On basis of this table, then, it can be concluded quite emphatically that  $S_{\text{pron}}$  was the preferred pronominal variant in the main clauses of the corpus texts – whether belonging to the early or the late period of the language.

#### 4.7.1.2 $S_\emptyset$ in second conjuncts

Similarly to section 4.7.1.1, the present section will offer a detailed presentation of the distribution of  $S_\emptyset$  and  $S_{\text{pron}}$  in second conjunct clauses according to texts, supplementing the general overview provided in table 4.6. Three clauses of this variety containing  $S_\emptyset$  are exemplified in (4.21)–(4.23).

- (4.21) *Þæs on þæm æfterran geare Hannibal sende sciphere on Rome,*  
 this-G in the following year Hannibal sent-SG ship-army to Rome  
 & [ $S_\emptyset$ ] *þær ungemetlice gehergeadon.*  
 and [they] there excessively ravaged-PL  
 ‘In the year after this, Hannibal sent a fleet to Rome and they there excessively ravaged.’  
 (Or 96.16)
- (4.22) *Þa gelamp hit þa ða hi on ðære byrig betleem wicodon. þæt hire tima*  
 then happened it then when they in the town Bethlehem sojourned that her time  
*was gefylled þæt hio cynnan sceolde: & [ $S_\emptyset$ ] acende þa hire frumcynnegan sunu.*  
 was come that she give-birth should. and [she] gave-birth-to then her firstborn son  
 ‘It happened, then, when they sojourned in the town of Bethlehem, that the time was  
 come for her to give birth, and she gave birth to her firstborn son.’  
 (ÆCHom I 190.16)
- (4.23) *Þa ætywde him Helias mid Moyses*  
 then appeared-SG to-them Elias with Moses  
 & [ $S_\emptyset$ ] *to him spræcon.*  
 and [they] to him spoke  
 ‘Then Elias and Moses appeared to them, and they were speaking to him [i.e. Jesus].’  
 (WSCp Mk 9.4)

Table 4.8 presents the distribution of  $S_{\text{pron}}$  and  $S_\emptyset$  in second conjunct clauses according to the corpus texts.

Table 4.8: S<sub>pron</sub> vs. S<sub>ø</sub> in second conjunct clauses according to texts

Texts	S <sub>pron</sub>		S <sub>ø</sub>		Total	
	n	%	n	%	n	%
Bede	584	93.6	40	6.4	624	100.0
Bo	333	97.9	7	2.1	340	100.0
ByrM	44	100.0	0	0.0	44	100.0
ChronA	131	94.9	7	5.1	138	100.0
CP	450	97.6	11	2.4	461	100.0
GD	1014	99.2	8	0.8	1022	100.0
LawAf	4	100.0	0	0.0	4	100.0
Or	365	93.1	27	6.9	392	100.0
WSCp	1135	99.0	11	1.0	1146	100.0
ÆCHom I	860	98.5	13	1.5	873	100.0
ÆCHom II	717	98.0	15	2.0	732	100.0
<b>Overall</b>	<b>5 637</b>	<b>97.6</b>	<b>139</b>	<b>2.4</b>	<b>5 776</b>	<b>100.0</b>

In combination with table 4.7, this table shows that S<sub>ø</sub> is more widely distributed in second conjuncts than in main clauses, demonstrating an overall frequency of 2.4%. This is a substantial increase compared to the overall frequency of 0.5% S<sub>ø</sub> observed in main clauses. The difference between non-conjunct and second conjunct clauses is statistically significant ( $\chi^2=117.37$ ,  $p\leq.0001$ ). Also, as seen in table 4.6, the distribution of S<sub>ø</sub> in second conjuncts is higher than in any of the clause types discussed here. The difference between second conjunct clauses and all the other clause types combined is also statistically significant ( $\chi^2=194.66$ ,  $p\leq.0001$ ). Additionally, it is observed that many more texts display relatively high frequencies in this clause type compared to the other clause types. Also, nine of the eleven texts feature S<sub>ø</sub> in this category. As the table shows, the frequencies range from 0.8% to 6.9% of the total, but “high” frequencies are not restricted to a single source, as the case was in table 4.7. The highest frequency is observed in *Orosius*, which features an entire 6.9% S<sub>ø</sub> in second conjunct clauses. *Bede* is very close behind with a frequency of 6.4%. The *Anglo-Saxon Chronicle* also displays what must be referred to as a notably high frequency, at 5.1%. At this point, it may be surmised that it is the relatively high frequency of empty subjects in second conjuncts that leads to the claims concerning the idiomatic nature of subject omission in OE, as discussed at several stages previously. This category is certainly the only one featuring statistics that even marginally support the – possibly impressionistic – claims made by many scholars of OE about the permissibility of structures featuring empty subjects.



### 4.7.1.3 S<sub>ø</sub> in relative clauses

Attention is now directed to the distribution of S<sub>ø</sub> in relative clauses. As clarified above, the few instances discussed here comprise empty pronominal subjects of relative clauses introduced by an object relative, as exemplified in (4.24) and (4.25).

- (4.24) *bæd he þæt heo him biscop onsende, þæs lare & þegnunge Ongolþeode,*  
asked he that they him bishop send, by-whose teaching and ministry English-people  
*þe [S<sub>ø</sub>] rehte, þæs Drihtenlecan geleafan gife leornade*  
which [he] ruled, the Lord's faith grace learnt  
'He begged them to send him a bishop, by whose teaching and ministry the English  
people, which he ruled, might learn the grace of God's faith.'  
(Bede 158.6)
- (4.25) *& þa þa he ham ne com in ða tid, þe [S<sub>ø</sub>] him bebead,*  
and then when he home not came in the time, which [he] him commanded  
*se Godes wer Florentius þa wende his hamcymes & his abad oþ æfentid*  
the God's man Florentius then expected his return and of-him waited until evening-time  
'And when he did not come home at the time which he had appointed to him, the man  
of God, Florentius, expected his return and waited for him until evening.'  
(GD (C) 207.1)

Only very few of the instances of S<sub>ø</sub> investigated here are located in relative clauses. The four instances classified as belonging to this category constitute only 1.5% of the total observed citations featuring empty subjects. Accordingly, as table 4.9 demonstrates, it may be stated quite unequivocally that the general distribution of S<sub>ø</sub> in relative clauses is extremely restricted, with tokens being observed in only two of the texts, namely *Bede* and Gregory's *Dialogues*. In addition, of the four tokens observed in total, three occur in *Bede*. Once again, then, it is observed that *Bede* provides examples of rarely observed variants of a phenomenon which is already exceedingly rare. Table 4.9 shows the distribution of S<sub>ø</sub> and S<sub>pron</sub> in relative clauses.

Table 4.9:  $S_{\text{pron}}$  vs.  $S_{\emptyset}$  in relative clauses according to texts

Texts	$S_{\text{pron}}$		$S_{\emptyset}$		Total	
	n	%	n	%	n	%
Bede	473	99.4	3	0.6	476	100.0
Bo	331	100.0	0	0.0	331	100.0
ByrM	15	100.0	0	0.0	15	100.0
ChronA	44	100.0	0	0.0	44	100.0
CP	346	100.0	0	0.0	346	100.0
GD	470	99.8	1	0.2	471	100.0
LawAf	9	100.0	0	0.0	9	100.0
Or	197	100.0	0	0.0	197	100.0
WSCp	209	100.0	0	0.0	209	100.0
ÆCHom I	321	100.0	0	0.0	321	100.0
ÆCHom II	364	100.0	0	0.0	364	100.0
<b>Overall</b>	<b>2 779</b>	<b>99.9</b>	<b>4</b>	<b>0.1</b>	<b>2 783</b>	<b>100.0</b>

As the table shows, the relative frequency of  $S_{\emptyset}$  in relative clauses is a measly 0.1%. This is without comparison the lowest frequency of  $S_{\emptyset}$  according to clause type seen in the entire study. Given the figure of 0.1%, it seems unnecessary to offer further comment, except to say that use of empty subjects in relative clauses does not appear to have been a viable possibility for the authors and scribes of the texts analysed here. However, a methodological problem was referenced in section 3.6.2 which is of particular relevance here, namely that it is very difficult indeed to achieve certainty regarding whether some of the tokens of  $S_{\emptyset}$  actually represent scribal errors. Since there are so very few tokens belonging to this category, it may very well be that some – or even all – of the instances grouped here do not represent grammatical OE, but rather mistakes by authors or scribes.

#### 4.7.1.4 $S_{\emptyset}$ in adverbial clauses

The following section will present the instances of  $S_{\emptyset}$  occurring in the adverbial clauses of the corpus texts. Examples (4.26) and (4.27) demonstrate adverbial clauses featuring empty subjects.

- (4.26) *Hu Boetius sæde hu swytole [ $S_{\emptyset}$ ] ongiten hæfde þæt hit eall soð wære*  
 how Boethius said how clearly [he] understood had that it all true was  
 ‘How Boethius said how clearly he understood that it was all true.’  
 (Bo (heading) 10.13)

- (4.27) *On þysum life we ateoriað gif [S<sub>ø</sub>] us mid bigleofan ne ferciað:*  
 in this life we waste-away if [we] us with with food not sustain  
 ‘In this life, we waste away if we do not sustain ourselves with food.’  
 (ÆCHom I 457.205)

Table 4.10 provides an overview of the distribution of S<sub>ø</sub> and S<sub>pron</sub> in adverbial clauses according to the corpus texts.

Table 4.10: S<sub>pron</sub> vs. S<sub>ø</sub> in adverbial clauses according to texts

Texts	S <sub>pron</sub>		S <sub>ø</sub>		Total	
	n	%	n	%	n	%
Bede	636	97.7	16	2.5	652	100.0
Bo	739	99.1	7	0.9	746	100.0
ByrM	50	100.0	0	0.0	50	100.0
ChronA	85	100.0	0	0.0	85	100.0
CP	1216	99.4	7	0.6	1223	100.0
GD	888	100.0	0	0.0	888	100.0
LawAf	51	100.0	0	0.0	51	100.0
Or	584	99.7	2	0.3	586	100.0
WSCp	947	99.6	4	0.4	951	100.0
ÆCHom I	1154	99.8	2	0.2	1156	100.0
ÆCHom II	871	99.9	1	0.1	872	100.0
<b>Overall</b>	<b>7221</b>	<b>99.5</b>	<b>39</b>	<b>0.5</b>	<b>7260</b>	<b>100.0</b>

The overall figure of 0.5% S<sub>ø</sub> in adverbial clauses is identical to the frequency seen in non-conjunct main clauses. This frequency is very low. Similarly to comments made regarding the situation of subject omission in main clauses, which displayed the same overall relative frequency, it must be concluded that overt pronominal subjects were the norm in this clause type, and that empty realisations of subject pronouns are only erratically observed. It is interesting, however, to note that *Bede* demonstrates a quite high frequency – relatively speaking – for S<sub>ø</sub> in this clause type. The 2.5% S<sub>ø</sub> boasted by this text is much higher than any of the other texts, the closest “contender” being the 0.9% of *Boethius*. The difference between *Bede* and the other texts combined is statistically significant ( $\chi^2=45.4$ ,  $p\leq.0001$ ).

#### 4.7.1.5 S<sub>ø</sub> in *þæt*-clauses

S<sub>ø</sub> is also observed in *þæt*-clauses in the texts under analysis here. Two examples of empty subjects occurring in this clause type have been given in (4.28) and (4.29).

(4.28) *Toæteacte eac swelce se gesið his benum, þæt he his tearas geat & weop & geornlice*  
 added also thus the gesith his entreaties, that he his tears shed and wept and earnestly  
*bæd & halsade, þæt he to ðæm untruman men ineode & him fore gebæde.*  
 prayed and entreated, that he to the sick man in-went and him for prayed  
*& sægde þæt [S<sub>ø</sub>] him leof wære & his lif niedbehæfdlic:*  
 and said that [he] him dear was-SBJCT and his life necessary  
 ‘The gesith also added his entreaties, and shed tears and wept and besought him to  
 visit the sick man and pray for him, and said that he was dear to him, and his life  
 necessary.’  
 (Bede 396.23)

(4.29) *Oft eac gebyreð ðonne se scrift ongit ðæs costunga ðe he him ondetted*  
 often also happens when the confessor hears-of the temptations which he to-him confesses  
*ðæt [S<sub>ø</sub>] eac self bið mid ðæm ilcum gecostod.*  
 that [he] also self is by the same tempted  
 ‘It happens often when the confessor hears of the temptations which he confesses to  
 him, that he himself is tempted by the same thing.’  
 (CP 105.19)

Again, a comparison has been provided between S<sub>ø</sub> and S<sub>pron</sub> as they appear in *þæt*-clauses in the corpus of texts. These statistics are presented in table 4.11.

Table 4.11: S<sub>pron</sub> vs. S<sub>ø</sub> in *þæt*-clauses according to texts

Texts	S <sub>pron</sub>		S <sub>ø</sub>		Total	
	n	%	n	%	n	%
Bede	548	97.2	15	2.7	563	100.0
Bo	497	99.4	3	0.6	500	100.0
ByrM	39	100.0	0	0.0	39	100.0
ChronA	23	100.0	0	0.0	23	100.0
CP	826	99.2	7	0.8	833	100.0
GD	821	99.8	2	0.2	823	100.0
LawAf	7	100.0	0	0.0	7	100.0
Or	321	99.4	2	0.6	323	100.0
WSCp	399	100.0	0	0.0	399	100.0
ÆCHom I	480	99.2	4	0.8	484	100.0
ÆCHom II	453	100.0	0	0.0	453	100.0
<b>Overall</b>	<b>4 414</b>	<b>99.3</b>	<b>33</b>	<b>0.7</b>	<b>4 447</b>	<b>100.0</b>

The overall relative frequency of 0.7% S<sub>ø</sub> in *þæt*-clauses is only marginally higher than the frequencies observed for main clauses and adverbial clauses (0.5%). This means that many of the conclusions drawn in relation with non-conjunct main, relative and adverbial clauses apply here, as well: it is very difficult to speak of a living idiom when clauses featured overt pronominal subjects in over 99% of the cases. However, attention should once more be drawn to the fact that while overall frequencies are low for most texts, *Bede* again displays a

relatively high frequency. The 2.8% observed here is considerably higher than the second-highest frequency – namely the 0.8% seen in CP and *ÆCHom I*. Again, the difference between *Bede* and the rest of the corpus texts is statistically significant ( $\chi^2=29.42$ ,  $p\leq.0001$ ).

#### 4.7.1.6 Summary

As has been demonstrated, then, there may be good cause to claim that the statements some scholars have made about empty pronominal subjects in OE have been somewhat exaggerated. The data provided in the preceding sections have shown that the distribution of  $S_\emptyset$  is actually very limited in all clause types except for second conjunct clauses, and even in this clause type, the frequency is very low. It has also been noted at several stages that *Bede* stands out in many of these categories. This text has consistently featured relatively high frequencies for this restricted phenomenon, even when the frequencies demonstrated by other texts have been negligible. In non-conjunct main clauses, for instance, percentages did not exceed 0.4% for any of the other texts, yet *Bede* featured  $S_\emptyset$  in 3.4% of the total cases where a clause had a pronominal subject. In this context of generally very low frequencies, such a frequency must be said to be comparatively high. It was also demonstrated for most of the categories that the difference between *Bede* and the rest of the corpus of texts was statistically significant.

The high overall relative frequency observed for second conjunct clauses is not due only to influence from *Bede*, however, as this clause type genuinely did feature higher frequencies on a general basis. This led to the conclusion that second conjunct clauses, at an overall frequency of 2.4%  $S_\emptyset$ , is the only clause type even marginally supporting the claim that empty subject pronouns are a regular or idiomatic feature of OE.

Additionally, if the observed frequencies for  $S_\emptyset$  in non-conjunct and second conjunct main clauses are combined and compared to the combined frequencies of all the types of subordinate clause, it can be observed that the difference between these clause types is statistically significant ( $\chi^2=36.2$ ,  $p\leq.0001$ ). This could be used as evidence in favour of the notion that overt pronominal subjects may first have started appearing in subordinate clauses, before spreading to main clauses (cf. section 2.7.2), as the empty variant is exceedingly rare in this particular environment. This could indicate a longer period of “competition” between overt and empty subjects in this environment, reflected in the “near-eradication” of  $S_\emptyset$  in subordinate clauses. Of course, it should be kept in mind that while frequencies for  $S_\emptyset$  in main clauses are significantly higher, the frequencies are still not particularly high.

## 4.7.2 Syntactic function of the antecedent

Attention will now be turned to the relationship between the empty subject pronoun and the expression with which it shares reference – its antecedent.<sup>4</sup> As previously indicated (cf. sections 1.2 and 3.3), subject omission by ellipsis is permissible in PDE when the subject of the second member of the co-ordinated structure is omitted as a consequence of being identical to the immediately preceding subject. The same is the case in OE. Such unexpressed coordinated subjects are easily recoverable, and presents no interpretational challenge for the listener. However, OE seems to have operated with less stringent rules for omission through co-reference, and there are in evidence quite a wide variety of combinations between empty subjects and antecedents. The following sub-sections will present the various clause elements with which the empty pronoun can share reference.

### 4.7.2.1 $S_\emptyset$ co-referent with a preceding or following subject

In the corpus of analysed texts, numerous instances of empty subjects are observed to share reference with the subject of a preceding – or in very rare cases, following – clause structure. One such clause is exemplified in (4.30). In this case, the empty pronoun in the main clause shares reference with the personal pronoun *he* functioning as subject in the preceding subordinate clause.<sup>5</sup>

(4.30) *swa swa he leohtlice onslepte, buton ænigre gefelnisse sares [ $S_\emptyset$ ] þone gast onsende*  
just as-if he lightly in-slept, without any sense of-pain [he] the spirit on-sent  
'As if he had fallen lightly asleep, without any sense of pain he gave up the ghost.'  
(Bede 296.14)

Evidence collected suggests that empty subjects sharing reference with a subject of a preceding clause fall into four main categories according to the type of clause (main or subordinate) in which the empty pronoun and its antecedent are found. While section 4.7.3 will provide a slightly more in-depth overview of the clausal location of the antecedent, these categories can briefly be summarised as follows: empty subjects occurring in main clauses which are co-referent with the subject of a preceding subordinate clause, as was seen in (4.30) above; empty subjects occurring in subordinate clauses which are co-referent with the subject of a preceding main clause as exemplified in (4.31) below; empty subjects occurring in main clauses which are co-referent with the subject of a preceding main clause, but where the two main clauses are not conjoined, and the gap and the antecedent are separated by other

---

<sup>4</sup> This term will also cover the rare cases of cataphoric reference.

<sup>5</sup> In the following examples, antecedents are indicated by underlining.

intervening referents or other factors complicating the relationship between gap and antecedent (4.32); and empty subjects of subordinate clauses which are co-referent with the subject of a preceding subordinate clause (4.33).

- (4.31) *Mine gebroðra: gyrstandæg gemedmode ure drihten hine sylfne*  
 my brothers: yesterday vouchsafed our Lord him self  
*þæt [S<sub>ø</sub>] ðysne middaneard þurh soðre menniscnysse geneosode:*  
 that [he] this earth through true humanity came-to  
 ‘My brothers: yesterday our Lord himself vouchsafed that he came to this earth as a true human’  
 (ÆCHom I 205.195)

As indicated above, then, example (4.31) demonstrates an instance of S<sub>ø</sub> located in a subordinate clause sharing reference with the subject of a preceding main clause.

- (4.32) *Ond heo ða aspyredon þæt, & hwonan he wæs: [S<sub>ø</sub>] genoman þa his lichoman,*  
 and they then asked-about that, and from-where he was. [they] took-PL then his body  
 ‘and they inquired about that, and about where he came from. They then took his body [...]’  
 (Bede 90.33)

Example (4.32) illustrates how the subject of a main clause shares reference with the subject of a preceding main clause, but where the empty pronoun and its antecedent are separated by another referent – namely the masculine pronoun *he*, referring to the dead man.

- (4.33) *Þonne is þeaw þæs apostolican seðles, þonne heo biscopas halgiað, þæt [S<sub>ø</sub>] him*  
 then is custom of-the apostolic see, when they bishops hallow, that [they] him  
*bebodu sellað,*  
 directions give-PL  
 ‘Now, it is the custom of the apostolic see, when consecrating bishops, to give him instructions [...]’  
 (Bede 64.13)

Example (4.33) shows an empty subject of a *þæt*-clause co-referent with the subject of a preceding adverbial clause of time.

In addition to empty subjects sharing reference with another “regular” subject, there are also in evidence a few cases where S<sub>ø</sub> is co-referent with more “unusual” subjects. Among these are instances of S<sub>ø</sub> which are co-referent with so-called *oblique subjects* (cf. section 4.7.2.2), instances of S<sub>ø</sub> which share reference with only “part” of a preceding subject, and instances of S<sub>ø</sub> which are co-referent with other elements in addition to a preceding subject. Examples of S<sub>ø</sub> co-referent with the subject of a following main or subordinate clause are also

seen. However, the latter types are exceptionally rarely seen in the texts analysed in conjunction with this work.

Now, as said, there arguably are present several cases of  $S_{\emptyset}$  which share reference with only part of a preceding subject, frequently signalled by a change in number in the inflection of the verb. One such instance is presented in (4.34).

- (4.34) *On byssum ealande com upp se Godes þeow Agustinus & his geferan;*  
on this island came-SG up the God's servant Augustine and his companions.  
*was he feowertiga sum.*  
was he forty-G.PL some.  
*Noman hi eac swylce him wealstodas of Franclande mid,*  
took they also likewise them-D interpreters from France with,  
*swa him Sanctus Gregorius bebead.*  
as them Saint Gregory asked.  
*& [S<sub>∅</sub>] þa sende to Æþelbyrhte ærenddracan*  
and [he] then sent-SG to Æthelberht messenger  
'To this island came God's servant Augustine, and his companions. They were forty in all. They brought with them interpreters from France, as Saint Gregory had instructed them. And he then sent a messenger to Æthelberht.'  
(Bede (O) 58.4)

Here, the antecedent is *se Godes þeow Agustinus*, 'God's servant Augustine'. The antecedent is a third person NP functioning as part of the subject in its clause. The subject as a whole, which also includes the NP *his geferan* 'his companions', takes the singular verb *com* 'came', despite arguably being a plural element.<sup>6</sup> The fact that the plural pronoun *hi* 'they' is used in combination with the plural verb *noman* 'took' the next time the group is referenced supports this "plural" analysis. However, the verb combining with the empty pronoun, *sende* 'sent', is in the singular. This indicates that the empty subject must be correspondingly singular, and thus it seems that  $S_{\emptyset}$  refers to 'Augustine' alone, and not the group as a whole. This shows that the empty subject need not correspond to a "full" clause element, but that the antecedent actually may be only part of such an element.

Narratively, then, the author alternates between singular and plural pronouns in order to shift the focus from the group as a whole to the leader of the group, namely 'Augustine'. This actually makes (4.34) one of the cases where the verbal morphology is absolutely instrumental in determining the intended reference of the empty pronoun, as this is the means

---

<sup>6</sup> According to Mitchell (1985), it is not unusual that a combination of two or more entities functioning as subject combine with a singular verb. He states that even though two subjects linked by *and* often take a plural verb, there are also cases where such subjects take one in the singular. This is relatively common when the verb precedes the first subject, for instance, as the case is in (4.34) (Mitchell 1985: 15–16).



by which the author distinguishes between ‘Augustine and his companions’ and simply ‘Augustine’. Note that the verb only distinguishes number here, not person.

Also interesting to note, no doubt, is the fact that  $S_\phi$  in (4.34) is co-referent with a third person NP occurring at quite some distance, and is in fact not co-referent with the immediately preceding third person NP, *Sanctus Gregorius* ‘Saint Gregory’. This means that a certain amount of pragmatic inferencing is necessary to identify the antecedent correctly, and shows that it was quite possible for  $S_\phi$  to occur at some distance from its antecedent. To the modern reader, at least, it can in very many cases be difficult to identify the antecedent in cases similar to (4.34), as there is considerable ambiguity regarding the identity of the empty subject. It seems possible that such interpretational difficulties must have been present for the speakers of OE, as well. A somewhat related example is given below:

- (4.35) & *ealne þone here he het mid þæm scipum þonan wendan þe he ær to geþoht hæfde,*  
and all the army he commanded with the ships from-there turn which he before to  
thought had,  
& [ $S_\phi$ ]<sub>i</sub> *up comon æt Leptan þæm tune,* & [ $S_\phi$ ]<sub>j</sub> *hrædlice for to Cartaina,*  
and [they] up came-PL to Leptis the town, and [he] readily went-SG to Carthage  
‘and he ordered all the army with their ships to turn from that place, which he had  
thought of, and they came to the town of Leptis, and he quickly went to Carthage.’  
(Or 107.11)

This is an interesting example in several ways. First, one observes two separate instances of  $S_\phi$  in two consecutive clauses, which can be said to refer to different antecedents. In the first instance,  $S_\phi$  takes the plural verb *comon* ‘came’, indicating that the empty subject should also be plural, while the second instance takes the singular verb *for* ‘went’. A possible analysis is thus that the antecedent of the first instance of  $S_\phi$  comprises both the subject *he*, referring back to *Hannibale* ‘Hannibal’, mentioned in Or 107.9, and the direct object *ealne þone here* ‘all the army’. Both elements located in the preceding main clause. Arguably, the antecedent also comprises an object of a preposition which may or may not be part of the direct object, namely *mid þæm scipum* ‘with the ships’, in addition to the subject and the direct object. The second instance refers only to Hannibal, represented by *he* ‘he’.

However, although the various syntactic elements comprising the antecedents here are clearly distinct, it might be argued that they in one sense represent the same entity or entities. That is, it may be argued that the singular ‘he’ encompasses both ‘Hannibal’ and the army, as the singular verb in the latter instance could be explained by ‘the army and the ships’ being – in manner of speaking – subordinate to and “included” in the singular pronoun ‘he’. If so, the

latter instance of  $S_\emptyset$  would still include ‘the army’, yet be in the singular on account of being focussed on the most prominent figure. If this were the case, the variation between the singular and plural would be explained, and the problem of how the reader is supposed to identify the distinct antecedents of two instances of  $S_\emptyset$  occurring in very close proximity is also avoided. According to Mitchell (1985), the combination of a singular verb with “two or more singular subjects joined by *and*” is not unusual, particularly when the two subjects are “thought of as a unit” (Mitchell 1985: 14). While the antecedent in this case is not two subjects joined by *and*, but rather a subject, a direct object and an object of a preposition, the proposed explanation is certainly a possibility in this particular instance. Example (4.36) provides counterpoint to (4.35) in doing more or less the opposite.

- (4.36) *He þonan afor, & his fierd gelædde on an oþer fæstre land,*  
 he from-there went, and his army led to another fast land,  
*& [S<sub>∅</sub>] þær gewunedon oþ niht.*  
 and [they] there dwelled-PL until night  
 ‘He went from there, and led his army to another fast-land, and they dwelled there until night.’  
 (Or 46.35)

Here, the plural *gewunedon* ‘dwelled’ indicates that both *he*, i.e. ‘Leonidas’, and ‘his army’ are intended as the antecedent of the empty subject. As such, the antecedent in example (4.36) comprises the subject and direct object of *separate preceding main clauses*. The two elements are obviously not considered a singular unit here, as the plural verb indicates.

In addition to cases where both a subject and an object form the antecedent of an empty pronoun, there are also a few examples where a subject and an object of a preposition constitute the antecedent of an instance of  $S_\emptyset$  which occurs with a verb inflected for the plural. One such instance has been given below.

- (4.37) *Ða astrehte se riggo hine to eorðan mid eallum his gefeþum. swiðe forhtigende*  
 then prostrated the Riggo him-RFLX to earth with all his companions, greatly fearing  
*þæt hi his fandian dorston. and [S<sub>∅</sub>] gecyrdon to heora hlaforde forhtmode cyðende hu*  
 that they it-G try dared. And [they] returned-PL to their master fearfully saying how  
*hrædlice hi arasode wurdon;*  
 quickly they discovered were  
 ‘Then Riggo prostrated himself along with all his companions, fearing that they should dare try it. And they returned to their master, speaking fearfully of how quickly they were discovered.’  
 (ÆCHom II 99.244)

Regardless of the numeral status of the verbs with which  $S_{\emptyset}$  combines, though, instances of “complex antecedents” such as those given above have been observed with some frequency in the course of this work, and they occur in both main and subordinate clauses. This provides further evidence that  $S_{\emptyset}$  need not be co-referent with a “whole” clause element, and it is shown that it may be co-referent more than one element. These “complex” antecedents are reminiscent of the “split” and “partial” antecedents distinguished in Sigurðsson (1993: 252) (cf. also section 2.4).

As we have seen, then, the situation regarding  $S_{\emptyset}$  and its antecedent can be quite complex. Still, the examples of  $S_{\emptyset}$  analysed here display a clear tendency toward co-reference with a subject occurring in the surrounding discourse. An entire 56.7% of the empty subject pronouns in the corpus of texts used in this work share reference with a preceding or following subject, whether nominal or pronominal. If those tokens of  $S_{\emptyset}$  that share reference with a subject and an additional element are included, the figure increases to 60.8%. If instances co-referent with an oblique experiencer antecedent – to be discussed below – are also included, the figure increases further to 66.0%. These figures make it quite clear that the empty subjects analysed here by far most often share reference with a preceding – or in rare cases following – subject, whether in combination with another element or not. This will become particularly clear when contrasted with figures for the other antecedent types identified for this work.

#### **4.7.2.2 Oblique experiencer antecedents**

As previously mentioned, a few instances of  $S_{\emptyset}$  have been observed to share reference with oblique experiencer antecedents. These are elements in the dative or accusative which arguably are best analysed as subjects, despite being in non-nominative cases. While these arguments are not strictly subjects in the traditional sense – that is, nominative elements controlling verbal agreement – there are some who believe that analysing examples such as those given below as objects is somewhat misrepresentative. For the purposes of this work, these arguments will be referred to as *oblique subjects*.

Oblique subjects, or *oblique subject-like NPs* in the terminology of Barðdal (2000), are not necessarily common, but are by the OE period still observed with some frequency in

combination with so-called impersonal verbs.<sup>7</sup> Three examples of  $S_\theta$  sharing reference with oblique experiencer antecedents have been given below.

(4.38) & *Romanum* wearð micel ege from him,  
and the-Romans-D became-SG great fear of him,  
& [ $S_\theta$ ] Uecilius þone consul ongean hiene mid firde sendon,  
and [they] Uecilius the consul against him with army sent-PL  
'And the Romans feared him greatly, and (they) sent the consul Uecilius against him  
with an army.'  
(Or 114.16)

(4.39) *Ða ofhreow þam munuce þæs hreoflian mægenleaste.*  
then grieved the monk-D of-the leper's helplessness  
& [ $S_\theta$ ] bewand hine mid his cæppan.  
and [he] draped him with his cape  
'Then the monk grieved for the leper's helplessness, and (he) draped him with his  
cape.'  
(ÆCHom I 369.139)

(4.40) *Þa scamode þone biscop & [ $S_\theta$ ] nolde him þa his costunge geandettan.*  
then was-ashamed the bishop-A and [he] not-would him then his temptation confess  
'Then the bishop was ashamed, and (he) would not confess his temptation to him.'  
(GD 190.8)

In (4.38),  $S_\theta$  is co-referent with the dative *Romanum* 'the-Romans', while the dative *þam munuce* 'the monk' serves the same function in (4.39). As seen in (4.38), the plural 'dative subject' does not agree with the singular verb *wearð* 'became-SG'. This is because only the nominative can control verbal agreement. Thus, clauses with oblique subjects consistently feature third person singular verbs (Haugland 2007: 347). In (4.40), the oblique subject is in the accusative, since the verb *scamode* 'was-ashamed' invariably takes an accusative experiencer.

It should be pointed out here that while these dative and accusative arguments for our purposes have been interpreted as the syntactic subjects of their clauses, this is not an uncontroversial decision. There has been significant scholarly debate as to the classification of elements such as these, and the general consensus still seems to be that oblique arguments such as the above should be interpreted as objects. However, scholars such as Allen (1999), Barðdal (2000) and Haugland (2007), argue that these oblique structures share many of the syntactic features of subjects, and that a subject-analysis is more fruitful than an object-

<sup>7</sup> Impersonal verbs are verbs which do not take a subject, or lack a 'personal' or referential subject. An example from Modern English would be *me thinks*, where the verb *think* in this context is impersonal, as there is no nominative *I* present.

analysis. An oblique subject analysis has been shown to be viable for many of the Old Germanic languages, and Barðdal states that “[o]blique subject-like NPs [...] pass all (known) subjecthood tests of Old Scandinavian and Old English” (Barðdal 2000: 25). While her main focus is on Old Scandinavian (i.e. ON, in the terminology used in this work), she claims that oblique subject-like NPs function as syntactic subjects in languages with “a similar structure and a similar case system” (Barðdal 2000: 28). This includes OE.<sup>8</sup>

Haugland points out that it is “well established for a number of languages [...] that non-nominative NPs can be subjects in terms of structural behaviour” (Haugland 2007: 357). Subjecthood tests referred to by Haugland include *positional criteria* – i.e. that oblique subjects have the same positional distribution as “regular” nominative subject – and *subject raising*, where oblique subjects appear as the “raised” subject of infinitive aspectual verbs such as *onginnan* ‘begin’. It is also shown that oblique subjects can appear in control infinitives. Additionally, it is demonstrated that oblique experiencers can control what Allen calls *coordinate subject deletion*, a fact which also indicates that these arguments function as subjects (Haugland 2007: 358-370). On the semantic level, it is pointed out that “[t]here is no agent role, but one argument is typically human and plays the semantic role usually referred to as the *experiencer*, i.e. someone affected by the bodily state”. The agent role is typically assigned to the subject, yet that of experiencer can also be associated with subject function, as the subject would then be the entity passively experiencing the action or process denoted by the verb.

In this study, 14 cases of  $S_{\emptyset}$  co-referent with an oblique subject have been observed. This means that 5.2% of the observed citations featuring  $S_{\emptyset}$  have an oblique experiencer antecedent.

#### 4.7.2.3 $S_{\emptyset}$ co-referent with a preceding object

While an entire 66.0% of the citations containing  $S_{\emptyset}$  examined in this work are co-referent with a subject, a subject and another argument or an oblique subject, there are also in evidence numerous instances of  $S_{\emptyset}$  sharing reference with a preceding object – whether direct or indirect. In other words, the empty subject may be co-referent with a non-nominative element found in the preceding context, including elements in the accusative or dative.<sup>9</sup>

---

<sup>8</sup>However, many of Barðdal’s own examples of “Old English” actually represent Middle English.

<sup>9</sup>No instances of  $S_{\emptyset}$  co-referent with an object in a following main or subordinate clause have been identified in the corpus texts examined here.

Examples of  $S_o$  co-referent with objects occurring in the surrounding context have been given in (4.41)–(4.43).

- (4.41) *& he geseah ðær micle menigo monna æghwæðeres hades;*  
 and he saw there great crowd persons of-both genders  
*& [S<sub>o</sub>] wæron missenlicræ ylðo & getincge men.*  
 and [they] were various ages and ranks persons  
 ‘and he saw there a great crowd of persons of both genders, and they were of various ages and ranks.’  
 (Bede 38.7)
- (4.42) *Lædde mon his lichoman to Lindesfarena ea, & [S<sub>o</sub>] in broðra lictune wæs bebyrged.*  
 led one his body to Lindisfarne island, and [it] in brethren’s graveyard was buried.  
 ‘His body was led to the island of Lindisfarne, and it was buried in the brethren’s graveyard.’  
 (Bede 204.5)
- (4.43) *& mon geseah swelce hit wære an gylden hring on heofonum brædre þonne sunne;*  
 and one saw as it were a golden ring in heavens broader than sun  
*& [S<sub>o</sub>] wæs from þæm heofone bradiende niþer oþ þa eorþan,*  
 and [it] was from the heaven reaching down to the earth  
 ‘and one saw in heaven, as it were, a golden ring broader than the sun, and it reached down from heaven to earth.’  
 (Or 123.20)

In all of the examples above, an instance of  $S_o$  occurring in a second conjunct clause is co-referent with the direct object of the preceding main clause. This does not, of course, mean that all the instances in this category are of this variety. As was seen in the cases where  $S_o$  shares reference with a preceding subject, the empty pronoun could occur in virtually any clause type, and so could the antecedent. Examples (4.44)–(4.46) serve to illustrate some of the ways in which the clausal locations of the empty subject and non-nominative antecedents could combine.

- (4.44) *& þa ðy seofoþan dæge, swa him gehaten wæs, æfter þon þe [S<sub>o</sub>] his forðfore*  
 and then the seventh day, as him promised was, after when [he] his departure  
*getrymede mid onfongennesse þæs Drihtenlecan lichoman & bloddes, þætte seo halige*  
 fortified by reception of-the Lord’s body and blood, that the holy  
*sawl was onlesed from þæs lichoman hefignessum*  
 soul was released from the body’s burdens  
 ‘and on the seventh day, as was promised to him, after he had fortified his departure by receiving the Lord’s body and blood, the holy soul was released from the burdens of the body’  
 (Bede 266.34)

- (4.45) *Ɔa offpuhte Ɔæt Mariuse Ɔæm consule, Iuliuses eame,*  
 then displeased that Marius the consul, Julius' uncle,  
*Ɔæt mon ðæt gewin nolde him betæcan. & [S<sub>ø</sub>] bæd Ɔæt him mon sealde Ɔone*  
 that one the war not-would him entrust and [he] asked that him one gave the  
*seofoƆan consulatum, & eac Ɔæt gewin,*  
 seventh consulship and also the war  
 'Then it displeased Marius, the consul, uncle to Julius, that one would not entrust the  
 persecution of the war to him and he asked that one should give him the seventh  
 consulship, and also the war.'  
 (Or 124.19)
- (4.46) *Ac him ða upƆlican onwrigenesse & weorc somod wiðstodon,*  
 but him then celestial obstacles and work together prevented,  
*Ɔæt [S<sub>ø</sub>] no hwæðer ðissa gefremedæ ne ðurhteah.*  
 so-that [he] not either of-these achieved nor carried-out.  
 'But then celestial obstacles and work together prevented him, so that he did not  
 achieve nor carry these out.'  
 (Bede 410.1)

Example (4.44) illustrates an instance of S<sub>ø</sub> in a subordinate clause co-referent with the indirect object of the preceding subordinate clause. More specifically, the empty pronoun is located in an adverbial clause of time, and the antecedent is located in an adverbial clause of reason. Example (4.45) shows an empty pronoun in a second conjunct clause sharing reference with the indirect object of the preceding subordinate *Ɔæt*-clause, while (4.46) demonstrates an instance of S<sub>ø</sub> occurring in an adverbial clause which is co-referent with the dative direct object of the preceding second conjunct clause. These examples should serve well to illustrate the fact that the empty subject and its antecedent could occur in various clause types and combinations of clause types in cases where the antecedent functions as object in its clause. It should also be noted that, in the same way as the antecedent could be comprised of a subject and an additional element, non-nominative elements could also combine in the role of antecedent. There are several examples where the antecedent of an empty subject is comprised, for instance, of a direct or indirect object, plus an object of a preposition.

In terms of quantification, then, instances of S<sub>ø</sub> sharing reference with an object are relatively commonly observed, yet not nearly to the degree of those which share reference with another subject, an oblique subject or a subject in addition to another element. The empty pronouns analysed in this work are co-referent with a single object – whether direct or indirect – in 20.5% of the cases. When those cases that are co-referent with an object in addition to another element are added, the figure rises to 23.5%. It may be added that if dative subjects

are interpreted as objects, which perhaps most scholars are prone to do, the figure would reach 28.7%.

#### 4.7.2.4 $S_\emptyset$ co-referent with an object of a preposition

A minor group of antecedents for  $S_\emptyset$  include objects of prepositions. Examples illustrating such antecedents are given below.

- (4.47) *We bebeodað þam deoflum þe on ðisum anlicnyssum sticiað þæt hi ut faron. and ða we command the devils which in these idols dwell that they out go and the anlicnysse tocwyson. þæt ge magon swa tocnawan. þæt sunne and mona. ne sind on idols crush. so-that you may so know. that sun and moon not are in ðisum anlicnyssum. ac [ $S_\emptyset$ ] sind mid deoflum afyllede; these idols. but [they] are with devils filled*  
'We command the devils dwelling in these idols to flee and the idols to be destroyed, so that you may thus realise that the sun and the moon are not in these idols, but that they are filled with devils.'  
(ÆCHom II 286.239)
- (4.48) *Ne nimð hig nan man æt me ac [ $S_\emptyset$ ] læte hig fram me sylfum.<sup>10</sup> not takes it not-one man from me but [I] lay-down it from me self*  
'No man will take it from me, yet I will lay it down myself.'  
(WSCp Jn. 10.18)
- (4.49) *He geseah ða þæt hine ne mihte nan læce gehælan. and sette his swurdes ord togenes he saw then that him not might not-one leech heal. and set his sword's point against his innoðe. and feol him onuppon. þæt [ $S_\emptyset$ ] him ðurheode his stomach. and fell it-D upon. so-that [it] him through-pierced*  
'He realised then that no leech could heal him, and set the point of his sword against his stomach and fell upon it, so that it pierced him through.'  
(ÆCHom II 279.213)

In example (4.47), an instance of  $S_\emptyset$  is co-referent with *ðisum anlicnyssum* 'these idols', the object of a preposition functioning as adverbial in the clause. Similarly, the empty pronoun in example (4.48) shares reference with *me* 'me', which is the object of the preposition *æt* 'at'. This sequence of preposition and its object serves the function of adverbial in its clause. Note also example (4.49), where the antecedent, while analysed as the object of a preposition here, actually is the object of a *postposition*, incidentally in the dative case.

This form of antecedent of an empty subject is, as indicated above, very rarely seen. It is observed in only 10 cases, which amounts to a frequency of 3.7% of the total number of antecedents. If the cases where the antecedent consists of an object of a preposition along with

---

<sup>10</sup> Liuzza (2000:204) lists this example among the "errors in the transmission of the text" (i.e. in the translation from Latin) due to the omission of the subject pronoun. The Latin version has the overt pronoun *ego* 'I', here.



a subject or object are included, the figure increases to 6.3%. This is still a very low frequency.

#### 4.7.2.5 $S_{\emptyset}$ co-referent with a genitive

Yet another very minor category of antecedent is observed in cases where empty subject pronouns are co-referent with a genitive occurring in the surrounding context. Two examples are given as illustration below.

- (4.50) [...] *æfter þon þe he wisse þæt Ecgfriþes willa wæs þæs cyninges ond his leafnes wæs.*  
after he knew that Ecgfriþe's will-G was the king's and his permission was.  
*Forþon þe [ $S_{\emptyset}$ ] him þa londare forgeaf & sealde, þe he þæt mynster on getimbrade.*  
because [he] him the property granted and gave, which he the monastery on  
constructed  
'[...] after he had ascertained that this was the wish of king Ecgfriþ, and was with his permission, because he had granted and assigned to him the property on which he erected the monastery.'  
(Bede 314.12)

- (4.51) *Bis ungesælige gear & þæt godlease gen to dæge laðe wunað, ge fore fleame cyninga*  
this unhappy year and the goodless still to day odious remains, both for flight of-kings  
*from Cristes geleafan & [ $S_{\emptyset}$ ] eft to deofolgyldum cerdon, ge for wedenheortnisse þæs*  
from Christ's faith and [they] after to devil-idols returned-PL, and for madness of-the  
*leodhatan Bretta cyninges.*  
tyrannical Britons' king.  
'This unhappy and godless year remains odious still to this day, both because of the apostasy of the kings from the Christian faith and their return to heathen idols, and the madness of the tyrannical British kings.'  
(Bede 154.6)

In example (4.50), the empty subject shares reference with *Ecgfriþes [...] þæs cyninges* 'King Ecgfriþ-G', an element in the genitive, while the antecedent in example (4.51) is the genitive *cyninga* 'kings-G'. Very few instances of  $S_{\emptyset}$  have been analysed as having a genitive antecedent. Only four citations in the source texts have been classified as belonging to this category, a figure which corresponds to a slight 2.2% of the total number of antecedents. Of these four, three are found in *Bede* and one in *Boethius*.

#### 4.7.2.6 "Miscellaneous" antecedents

There are also in evidence a number of empty subject pronouns demonstrating more complex situations in terms of co-reference compared to the main categories outlined above. These are antecedents which for various reasons do not conform to any of these categories. Due to the non-uniformity of these antecedents, a 'miscellaneous' category was created to represent

them. The antecedents in this category have in common that they do not occur as clause-elements, but have other functions. Examples are given in (4.52)–(4.54)

(4.52) *Forðon se apostol Sanctus Paulus mid ðy [S<sub>ø</sub>] cwæð [...]*  
 for the apostle Saint Paul when [he] said  
*he ða sona se apostol underðeodde & æftercwæð [...]*  
 he then at-once the apostle added and after-said  
 ‘Because when the apostle, Saint Paul, said [...], he at once agreed and replied [...].’  
 (Bede 82.30)

(4.53) *Ða he ðis cwæð, sona in ða ilcan tid ða cristnade he mæc.*  
 when he this said, at-once in the same time then baptised he me.  
*Ða [S<sub>ø</sub>] wæs geworden, þæs ðe he on minre ondwlitan bleow,*  
 when [it] was happened, when he on my face blew,  
*ða sona instepe gefelde ic mec batiende & werpende.*  
 then at-once felt I me growing-better and recovering  
 ‘At the same time as he said this, he baptised me. When it was done, after he had blown on my face, I immediately felt myself recovering.’  
 (Bede 402.33)

(4.54) *Se ðonne ðe wilnað ðæt [S<sub>ø</sub>] wolde on ðæm angienne his lifes woroldspedig weorðan*  
 he then who desires that [he] would in the beginning of-his life wealthy become  
*mid unryhte, se hine wile selfne [...]*  
 with unright, he him will self  
 ‘He who desires to become wealthy with unrighteousness in the beginning of his life, will himself [...].’  
 (CP 333.1)

Example (4.52) demonstrates an instance where the antecedent of the empty subject pronoun is a left-dislocated element, *se apostol Sanctus Paulus* ‘the apostle Saint Paul’. This left-dislocated element, which can be described more or less as an anticipatory “copy” of the subject, stands outside the clause structure, and thus cannot be said to have a clausal “location” or even a real clausal “function”. Example (4.53) illustrates a case where the antecedent of the empty subject is the action or process indicated by a verb – *cristnade* ‘baptised’. Another quite noteworthy example is given in (4.54), where the empty subject can actually be said to be part of its own antecedent, which in turn is a left-dislocated subject.

Only a total of seven instances have been classified as belonging to the “miscellaneous” category. This corresponds to a mere 2.6% of the citations featuring an antecedent. Of these seven instances, four can be found in *Bede*. The remaining three are found in *Cura Pastoralis*. Most of the antecedents placed in this particular category correspond to the examples given above.

#### 4.7.2.7 “Indefinite” antecedents

Finally, some instances of  $S_{\emptyset}$  are not co-referent with a syntactic entity occurring in the surrounding context, but instead have indefinite or generic reference. In some of these cases, the indefinite pronoun *mon/man* ‘one’ can be inserted in the vacant subject position. This means that there is no syntactic co-reference involved, yet it may be possible to talk about “notional” co-reference, where the indefinite entity corresponding with  $S_{\emptyset}$  can be inferred from the context. Such indefinite antecedents are in some instances used as agent-suppressing devices, and thus correspond to cases where passive constructions would be used in PDE (Allen 1999: 57). One such instance is given in (4.55) below, along with other examples of  $S_{\emptyset}$  where no overt antecedents are involved. Notice that in addition to serving as agent-suppressing devices, indefinite antecedents could have been used to “down-tone” or “gloss over” catastrophic events, such as those recounted in examples (4.56) and (4.57). In both of these cases, the implied antecedent is *manega* ‘many’, and the idea expressed is that battles and epidemics will cause or have caused the deaths of a considerable number of people. Now, in principle, it could be possible that the intended antecedent is not ‘many’, but rather another, similarly indefinite pronoun, such as ‘some’. However, this does not correspond with the disastrous nature of the presented events. In support of this, Thorpe (1846) has inserted the word *manega* ‘many’ in brackets in the OE version of the text, along with the corresponding *many* in the translation (Thorpe 1846: 482-3). The use of  $S_{\emptyset}$  instead of a pronoun here is to my mind somewhat reminiscent of an act of censorship by the author or scribe, perhaps due to a consideration of the sensibilities of his readers. This may be speculation only, and impressionistic speculation at that, but it is certainly difficult to see another explanation for  $S_{\emptyset}$  in (4.56) and (4.57), as the implied indefinite pronoun does not function as an agent-suppressing device here. This could indicate that other concerns determine the “omission” of a subject in such instances, although this is by no means certain.

- (4.55) *He befran ða hwam ða gebytlu gemynte wæron. swa mærllice getimbode;*  
he asked then for-whom-D the buildings meant were. so gloriously constructed  
*Him wæs gesæd. þæt hi wæron gemynte anum sutere on romana byrig.*  
him-D was told that they were meant for-one shoemaker-D in Rome city  
*And [S<sub>∅</sub>] hine eac namode*  
and [one] him also named-SG  
‘He then asked for whom the buildings were meant, that were so gloriously  
constructed. He was told that they were meant for a certain shoemaker in Rome, and  
one also named him.’  
(ÆCHom II 203.121)

- (4.56) *Ða dydon hi þurh þæs ealdormannes bene þæt ða deoflu spræcon swa swa heora*  
 then did they through the alderman’s prayer that the devils spoke so as their  
*gewuna wæs. and sædon þæt þær wære micel gefeoht toward.*  
 wont was. and said that there was great fight at-hand.  
*and on ægðre healfe [S<sub>ø</sub>] sceoldon feallan;*  
 and on either side [many?] should-PL fall.  
 ‘They did then, at the alderman’s prayer, so that the devils spoke, as was their wont,  
 and said that a great fight was at hand, and on either side many should fall.’  
 (ÆCHom II 280.23)
- (4.57) *Mid þy seo hreonis þæs oft cwedenan wooles feor & wide all wæs forhergende &*  
 when the storm of-the often mentioned pestilence-G far and wide all was ruining and  
*forneomende, þa cwom he eac swylce in þone dæl þæs mynstre, þe ða wæpnedmen in*  
 wasting, then came it-M also likewise in the part of-the monastery, which the men in  
*wæron; ond [S<sub>ø</sub>] dæghwamllice gehwær of weorulde to Drihtne genumene wæron.*  
 were. and [many?] daily on-all-sides from world to Lord taken were.  
 ‘When the storm of the often-mentioned pestilence had ruined and wasted far and  
 wide, it also came to the part of the monastery where the men were. And many on all  
 sides were daily taken from this life, to the Lord.’  
 (Bede 282.25)
- (4.58) *Nu sculon [S<sub>ø</sub>] herigean heofonrices weard*  
 now should-PL [we] praise heaven’s guard  
 ‘We should now praise Heaven’s guard.’  
 (Bede 344.2)

The vast majority of the instances of S<sub>ø</sub> in this study have overt antecedents. Only twelve cases of indefinite antecedents have been observed, a figure which amounts to 4.5% of the total citations featuring S<sub>ø</sub>. As (4.56)–(4.58) show, indefinite antecedents may well combine with plural verbs, and in fact, four of these twelve cases represent instances of plural verbs combining with the empty pronoun. In example (4.58), from *Cædmon’s Hymn*, one of the instances featuring a plural verb, the empty subject has been interpreted as representing a generic *we*.

### 4.7.3 Location of the antecedent

As previously mentioned, the antecedent of the empty subject may be located in a wide variety of clausal locations. In this context, four main types of antecedent locations have been identified, namely preceding main clauses, following main clauses, preceding subordinate clauses and following subordinate clauses. A “miscellaneous” category corresponding with that discussed in section 4.7.2.5 has also been identified. Those instances which do not have an overt antecedent, but rather relies on generic or indefinite reference, as per section 4.7.2.6, have been placed under the “indefinite” label.

Some attention has already been given to the location of the antecedent in the sections dealing with the function of the antecedent. These topics are closely intertwined and may in many instances be difficult to separate effectively. Thus, this section will attempt not to reiterate the points given above, but instead seek to provide a succinct survey of the distribution of antecedents in the various clause types identified. Table 4.12 details the clausal distribution of the overt antecedents of the instances of  $S_\emptyset$  analysed in this work.

Table 4.12: Clausal location of the antecedent of  $S_\emptyset$

<b>Location of the antecedent</b>	<b>n</b>	<b>%</b>
Preceding main clause	139	54.3
Following main clause	2	0.8
Preceding subordinate clause	108	42.2
Following subordinate clause	1	0.4
Miscellaneous	6	2.3
<b>Total</b>	<b>256</b>	<b>100.0</b>

Table 4.12 shows that the antecedent is most frequently located in a preceding main or subordinate clause. The antecedent occurs in a preceding main clause in 54.3% of the cases where  $S_\emptyset$  has an overt antecedent, and in a preceding subordinate clause in 42.2% of the cases. The combined frequency for these two clause types is 96.5%. The antecedent is located in a following main or subordinate clause in 1.2% of the cases, while miscellaneous antecedents comprise the remaining 2.3%. These figures should make it abundantly clear that  $S_\emptyset$ 's regularly anaphoric, and only exceptionally cataphoric. The difference between preceding main and subordinate clauses is not statistically significant at the five percent level (chi-square goodness of fit,  $\chi^2=3.64$ ,  $p\leq 0.06$ ). However, it *is* statistically significant at the six percent level, and it should be kept in mind that the customary cut-off point at five percent is more or less arbitrary. It is thus difficult to determine how much value should be attributed to this result. As a consequence, it is also difficult to conclude whether the permissibility of  $S_\emptyset$  was determined by the position of the antecedent in a preceding main or subordinate clause.

#### **4.8 “Distance” between $S_\emptyset$ and antecedent**

Another factor which we may hypothesise to be relevant to the discussion of  $S_\emptyset$  in OE is the “distance” between the empty pronoun and its antecedent, i.e. whether  $S_\emptyset$  and antecedent are separated by longer stretches of text, or whether the two occur in relatively close proximity. As noted in Chapter 3, it has been decided to count the number of words separating the two. Table 4.13 presents the results of this analysis. It should be unnecessary to point out that the

statistics given below pertain only to the instances of  $S_{\emptyset}$  sharing reference with an overt antecedent, excluding those with indefinite reference. The data below are given according to texts, grouped in four categories with arbitrary cut-off points.

Table 4.13: The distance between  $S_{\emptyset}$  and its antecedent

<b>Texts</b>		<b>1-3</b>	<b>4-6</b>	<b>7-10</b>	<b>10+</b>	<b>Total</b>
Bede	n	20	24	23	38	105
	%	<b>19.0</b>	<b>22.9</b>	<b>21.9</b>	<b>36.2</b>	<b>100.0</b>
Bo	n	2	10	2	6	20
	%	<b>10.0</b>	<b>50.0</b>	<b>10.0</b>	<b>30.0</b>	<b>100.0</b>
ChronA	n	0	4	0	3	7
	%	<b>0.0</b>	<b>57.1</b>	<b>0.0</b>	<b>42.9</b>	<b>100.0</b>
CP	n	7	7	10	4	28
	%	<b>25.0</b>	<b>25.0</b>	<b>35.7</b>	<b>14.3</b>	<b>100.0</b>
GD	n	3	3	2	3	11
	%	<b>27.3</b>	<b>27.3</b>	<b>18.2</b>	<b>27.3</b>	<b>100.0</b>
Or	n	5	4	10	11	30
	%	<b>16.7</b>	<b>13.3</b>	<b>33.3</b>	<b>36.7</b>	<b>100.0</b>
WSCp	n	4	4	2	7	17
	%	<b>23.5</b>	<b>23.5</b>	<b>11.8</b>	<b>41.2</b>	<b>100.0</b>
ÆCHom I	n	6	9	5	1	21
	%	<b>28.6</b>	<b>42.9</b>	<b>23.8</b>	<b>4.8</b>	<b>100.0</b>
ÆCHom II	n	6	5	3	3	17
	%	<b>35.3</b>	<b>29.4</b>	<b>17.6</b>	<b>17.6</b>	<b>100.0</b>
<b>Total</b>	n	53	70	57	76	256
	%	<b>20.7</b>	<b>27.3</b>	<b>22.3</b>	<b>29.7</b>	<b>100.0</b>

As the table shows, antecedents are quite equally distributed in terms of the distance separating them from the empty subject pronoun. The overall frequencies are all in the range between 20.7% (1–3 words) and 29.7% (more than 10 words). It must be concluded that no discernable pattern emerges here, other than the fact that the distribution is notably even. It could also be observed here that the works of Ælfric generally operate with very few words separating  $S_{\emptyset}$  and its antecedent. In the first series of the *Homilies*, only one instance of  $S_{\emptyset}$  is separated from the antecedent by more than ten words. In the second series, the same is true for three instances. In both texts, the empty pronoun is located within six words of its antecedent in the majority of the cases – this being true in 71.5% of the cases in ÆCHom I and 64.7% of the cases in ÆCHom II. In this connection, it should be recalled that Ælfric is considered one of the great teachers of his age, and that his prose is renowned for its accessibility. Perhaps the relatively short distance between  $S_{\emptyset}$  and antecedent attested in his texts reflects this fact. Overall, however, the main conclusion to be drawn from table 4.13 is

that  $S_{\emptyset}$  does not necessarily occur close to its antecedent. This is perhaps most strongly demonstrated in *Bede*, as no less than half of the instances where more than ten words separate  $S_{\emptyset}$  and its antecedent are from this source. *Bede* is in parts considered to be a rather poor translation, often displaying very complex structures and generally “difficult” language. In terms of presentation and ease of interpretation, then, the venerable Bede can in many ways be said to be the polar opposite of Ælfric, and this is no doubt reflected in the fact that half the tokens in the 10+ category comes from this text. The higher text-internal rates in the 10+ category for ChronA and WSCp are based on very few tokens in all, and thus not comparable to the situation in *Bede*. *Orosius*, another text considered syntactically “difficult”, also features relatively many cases where  $S_{\emptyset}$  and the antecedent are separated by more than ten words.

It should perhaps also be clarified here that the 10+ category encompasses examples with considerably more than 11 words separating the empty subject and its antecedent. In some cases, the two are separated by quite considerable stretches of texts, and the empty pronoun may additionally be separated from its antecedent by several intervening pronouns, complicating the process of identification further. In fact, textual distances of around 20 words are not uncommon. In example (4.34), for instance,  $S_{\emptyset}$  and its antecedent are separated by 22 words. However, even longer distances have been observed, as illustrated in example (4.61) below, where the two are separated by more than 40 words. This topic will be treated further in the immediately following section.

#### **4.9 Identification of the antecedent and possible explanations for $S_{\emptyset}$**

The preceding sections have presented and discussed the data collected on  $S_{\emptyset}$  and its various types of antecedent. It has also been stated that it is not always a straightforward matter to identify the antecedent. This section will expand on this notion, and elaborate slightly on the identification of the antecedent of  $S_{\emptyset}$ , as this is no doubt a central concern in correctly deciphering the meaning of a clause containing an empty pronoun. Consequently, the ease or difficulty in identifying the antecedent may contribute to explaining why this phenomenon appears in OE.

As seen in Chapter 2, a central concept in the generative studies of empty pronominal subjects is that verbal inflections serve to unambiguously make clear the referent of the empty pronoun. As stated in section 2.3, this is referred to as the “identification hypothesis”. When confronted with data from languages such as Vietnamese, which does not feature a verbal morphology yet still has empty pronominal subjects, newer generative studies have

introduced addenda to existing *pro*-drop theories intended to explain the discrepancies – but central importance is still placed on the role of verbal inflections. Now, while the focus on verbal inflections is certainly valid for many archetypical *pro*-drop languages, such as Spanish and Italian, it is doubtful whether this hypothesis is able to fully explain  $S_0$  in OE. By the OE period, the Germanic inflectional system had “deteriorated” to the point where the verbal morphology was in some cases insufficiently able to distinguish person unambiguously. As a case in point, tables 4.14 and 4.15 illustrate the various inflections of two common OE verbs, one strong (table 4.14) and one weak (table 4.15).

Table 4.14. Example of OE verbal inflections: *helpan* ‘help’ (strong, class 3b verb)

	Present		Preterite	
	Indicative	Subjunctive	Indicative	Subjunctive
<b>1 sg</b>	helpe	helpe	healp	hulpe
<b>2 sg</b>	hilpst	helpe	hulpe	hulpe
<b>3 sg</b>	hilpþ	helpe	healp	hulpe
<b>Pl</b>	helpaþ	helpen	hulpon	hulpen

Table 4.15. Example of OE verbal inflections: *hælan* ‘heal’ (weak, class 1 verb)

	Present		Preterite	
	Indicative	Subjunctive	Indicative	Subjunctive
<b>1 sg</b>	hæle	hæle	hælde	hælde
<b>2 sg</b>	hælst	hæle	hældest	hælde
<b>3 sg</b>	hælp	hæle	hælde	hælde
<b>Pl</b>	hælaþ	hælen	hældon	hælden

As is evident from the tables, the verbal inflections of OE are still quite complex, especially in comparison with the modern language. It is observed that number is distinguished in all cases, except in the subjunctive.<sup>11</sup> It is also seen that person is generally distinguished in the present indicative singular, leaving no doubt as to whether the referent is a first, second or third person entity. In the past indicative singular, however, more syncretism is present. For both the strong and the weak verb, no inflectional distinction is made between the first and third persons. Thus, situations may arise when a verb featuring a ZERO ending – illustrated by the strong verb – or an *-e* ending – illustrated by the weak verb – may be ambiguously first or third person.

The failure of the OE verbal system to distinguish between the first and third person in the past indicative singular certainly seems to be a crucial flaw when relying on the

<sup>11</sup> It should be noted that person is not distinguished at all in the subjunctive, but this, as Haugland (2007) points out, is a feature “OE shares with a prototypical *pro*-drop language like Italian” (Haugland 2007: 59).



morphological system to identify the antecedent of an empty subject. This problem comes to the fore particularly in narratives relying on the past tense to relay information, which all the corpus texts do to a great extent. This is perhaps particularly true for texts such as *Bede* and *Orosius*, which are histories, but the past tense is generally employed much more frequently than the present, which seems to have been used chiefly in cases where direct speech is delivered by one of the characters. In this way, the “weaknesses” of the verbal inflections in the past tense become more prominent for our purposes, as the majority of the verbs in our citations are in the past tense.

Now, this should be nuanced by pointing out that the lack of distinction between the first and third persons in the past indicative singular may not be such an immense problem, as the third person is far more common than the first in the corpus texts – and indeed quite possibly in most narrative genres. The first person seems to be used primarily in direct speech, as exemplified in (4.59)–(4.60).

(4.59) *ða cwæð se Wisdom: [...] ac ic wat hu [S<sub>ø</sub>] þin man tilian sceal.*  
 then said the Wisdom: [...] but I know how [I] your evil cure shall  
 ‘then wisdom said: [...] but I know how I can cure your evil.’  
 (Bo 13.15)

(4.60) *Ða ondsworede ic him: Ic hæbbu, cwæð [S<sub>ø</sub>], þæs treowes dæl, þe his heafod on*  
 then answered I him: I have, said [I], the tree’s part, which his head on  
*aseted wæs, þe he ofslegen wæs.*  
 placed was, when he killed was  
 ‘Then I answered him: I have, said I, the part of the tree on which his head was put,  
 when he was killed.’  
 (Bede 192.6)

At this point, then, it may be interesting to provide a brief overview of S<sub>ø</sub> with regard to the person of the pronoun to which the empty subject pronoun corresponds. This information is given in table 4.16.

Table 4.16. Distribution of S<sub>ø</sub> according to person

Person	S <sub>ø</sub> n	%
1st	7	2.6
2nd	8	3.0
3rd	252	94.4
<b>Total</b>	<b>267</b>	<b>100.0</b>

As the table shows, of all the analysed citations featuring  $S_{\emptyset}$ , the empty subject was interpreted to correspond to a first person pronoun in only 2.6% of the cases. Second person pronouns occur slightly more frequently, at 3%. In comparison,  $S_{\emptyset}$  corresponded to a third person pronoun in 94.4% of cases. On this basis, it may perhaps be claimed that a third person antecedent is more likely in many of the cases where verbal inflections ambiguously indicate either the first or the third person – yet this is a problematic conclusion. The distribution of first, second and third person pronouns, both overt and empty, probably depends on the type of text, and without data contrasting the distribution of overt and empty subjects with all persons, no solid conclusion can be reached. Van Gelderen (2000: 133, 149), however, has provided contrastive data of this kind, showing that  $S_{\emptyset}$  in fact is more frequent in the third person.

Additionally, however, inflectional syncretism was more extensive than tables 4.14 and 4.15 would indicate, which means that distinctions between person and possibly even number could be “muddled” even further (cf. also Haugland 2007: 59–60). This point offers good occasion to reiterate Mitchell’s statement that “the verb endings were too ambiguous [...] to have played much part” in the identification of  $S_{\emptyset}$  (Mitchell 1985: 628). However, it must be reiterated and commented specifically here that the verbal inflections are fully capable of distinguishing between the singular and the plural. In fact, as we have seen, the verb is in some cases instrumental in distinguishing between singular and plural antecedents. A good example of this was given in (4.34), where the singular inflection of the verb is the only thing indicating that  $S_{\emptyset}$  refers to ‘Augustine’ alone, as opposed to ‘Augustine and his companions’. This means that, while there is a great deal of syncretism in the OE inflectional morphology, verbal inflections could still play a disambiguating role in the case of number.

We are left with a situation, then, where the verbal system generally is able to distinguish between the singular and the plural, but is only variably able to distinguish between person. Given this, it must be concluded – in full agreement with Mitchell – that while the verbal inflections of OE in some instances are sufficient and helpful in the identification of the antecedent, they were in many cases not “detailed” enough to perform this task in a satisfying manner. This is contrary to the views promulgated by the *pro*-drop theories presented in Chapter 2, and must also be said to be directly contrary to van Gelderen’s claim that *pro*-drop is “common” in OE as a result of “the strength of the verbal person features” (van Gelderen 2000: 121). Indeed, van Gelderen claims that “the third person

is in fact more specified [than the first and second person, in terms of phi-features<sup>12</sup> – KAR ]” (van Gelderen 2000: 132), which leads her to concluding that third person pronouns may be dropped more easily (van Gelderen 2000: 194). This may very well be the case, yet, it seems she fails to consider that third person inflections in the past indicative are far less detailed than those for the present, and the fact that the preterite is used far more frequently than the present in the OE texts analysed here effectively serves to undermine van Gelderen’s claims about the influence of strong inflectional third person features. While her contrastive data does show that  $S_{\emptyset}$  is more frequent in the third person than the first and second, her conclusions regarding third person verbal inflections is in my view unfounded. Given that the third person was essentially indistinguishable from the first person in the preterite, and also provided that the majority of OE prose texts are in the preterite, it does not seem likely that the verbal inflections constitute an “absolute” key to explain the overweight of  $S_{\emptyset}$  corresponding to third person pronouns. There is also the fact to be considered that van Gelderen bases her analysis on data from the Lindisfarne and Rushworth gospels, which are interlinear glosses, and thus not representative of idiomatic OE.

Still, the fact that third person pronouns are more frequently left empty is probably not affected by this methodological oversight, and remains unexplained. As recalled, in conjunction with OHG (section 2.7.2), it was seen that  $S_{\emptyset}$  was more common than  $S_{\text{pron}}$  in the third person only. This observation corresponds with that of van Gelderen (2000). However, unlike van Gelderen, Axel (2005: 35) does not attribute this to the strength of the third person verbal features. The question, then, is whether there is anything remaining to explain this overweight of empty third person pronouns if we reject van Gelderen’s hypothesis.

It may be that there are properties of third person pronouns that make omitting them “easier” – or, alternatively, there are properties associated with first and second person pronouns that make omitting them “harder”. It is, however, doubtful whether these properties are of a syntactic or morphological nature, and perhaps pragmatic factors may be better suited to explain this uneven distribution. No such explanation will be attempted here, but it may be idly speculated that first and second person pronouns may be more important in identifying the actors in a text, as pronouns such as *ic* ‘I’ and *eow* ‘you’ are more specifically linked to human identity than the more general *he* ‘he’ and *seo* ‘she’.

In terms of inflectional ambiguity, the instances of  $S_{\emptyset}$  analysed in this work make clear that empty pronouns were not restricted to appearing only in contexts where verbal inflections

---

<sup>12</sup> *Phi-feature*, frequently represented simply as  $\emptyset$ , is a generative term indicating the information on person, number and gender contained in a noun or pronoun.

indicated person and number – as would be expected if it were true that verbal morphology features determine this phenomenon. My data has, in fact, shown that  $S_\emptyset$  frequently co-occurs with ambiguous verbs. There are several examples of empty pronouns combining with subjunctive verbs which, as recalled, are incapable of distinguishing person. Also, while most verbs distinguish effectively between the singular and the plural, plural verbs are, like subjunctives, incapable of distinguishing person. There is only a single inflection for the plural, both in the present and in the preterite. Accordingly, verbs inflected for the plural could be considered ambiguous in terms of identifying the antecedent. The fact that 34.5% of the instances of  $S_\emptyset$  combine with a plural verb should be indicative of the fact that the verbal morphology in many cases was insufficient in identifying the referent of the empty subject. Thus, it must be concluded that while verbal inflections can disambiguate the referent of an empty subject in some cases, particularly in differentiating between singular and plural referents, it does do so systematically. It follows, then, that the “identification hypothesis” has not been corroborated.

Coupled with inflectional ambiguity is the fact that, as seen for instance in example (4.34),  $S_\emptyset$  may not necessarily be close to the last iteration of its antecedent. As also mentioned in section 4.8,  $S_\emptyset$  and the antecedent could well be separated by several intervening referents. Now, if it is true that pronominal subjects are left empty on basis of being “redundant” because they are easily identifiable due to grammatical or contextual factors, it would be expected that the empty pronoun would share the grammatical function of its antecedent, and also be “close” to it in terms of textual distance. As we have seen, this is not necessarily the case, as  $S_\emptyset$  often is co-referent with non-nominative elements or parts of elements, and are also often quite distant from the antecedent. Additionally, the citations often display confusing alternation between various distinct referents, which makes correct identification of the empty subject even more difficult. For instance, cases have been observed where several occurrences of *he* ‘he’ intermingle without further identifying markers in the discourse surrounding an instance of  $S_\emptyset$  which correspond to a third person masculine antecedent. Consider example (4.61), where the different referents are identified by subscript letters in the glosses.

(4.61) *Ɔa eode he inn, swa swa he his hlafordes ærendo secgan scolde.*  
 then walked he<sub>i</sub> in, as if he<sub>i</sub> his lord's errand say should  
*Ond mid þy he þa geswippe muþe licetende ærend wreahte & lease fleosewade,*  
 and when he<sub>i</sub> devious mouth-D feigned errand reported and falsely whispered  
*þa astod he semninga, & getogene þy wæpne under his sceate, rædde on þone cyning.*  
 then rose he<sub>i</sub> suddenly and drawing the weapon under his garment, rushed on the king  
*Ɔa þæt þa Lilla geseah, se cyninges þegn him se holdesta, næfde he scyld æt honda,*  
 when that then Lilla<sub>j</sub> saw, the king's servant him the firmest, not-had he<sub>j</sub> shield at hand  
*þæt he þone cyning mid scyldan meahhte: [Sø]<sub>j</sub> sette þa his lichoman betweoh beforan*  
 that he<sub>j</sub> the king with shield might. [he]<sub>j</sub> placed then his body between before  
*þam styngre. & [Sø]<sub>i</sub> þurhstong þone cyninges þegn & þone cyning gewundade.*  
 the thrust, and [he]<sub>i</sub> through-struck the king's servant and the king wounded  
 'then he walked in, as if he would deliver his lord's errand. And as he then, with  
 devious words, reported his feigned errand and whispered falsely, he then rose  
 suddenly and, drawing the weapon under his garment, rushed on the king. When Lilla  
 saw this, who was the most devoted of the king's servants, having no shield at hand to  
 defend the king, he interposed his body to meet the thrust. And he pierced the king's  
 servant and wounded the king.'  
 (Bede 122.21)

Here we actually have two instances of S<sub>ø</sub>, which refer to distinct and separate third person masculine antecedents, namely 'Eomer' (he<sub>i</sub>, mentioned<sub>j</sub> by name in 122.8) and 'Lilla' (he<sub>j</sub>). It is observed that in this case, the distance between S<sub>ø<sub>i</sub></sub> and he<sub>i</sub> is particularly long, as the two are separated by more than 40 words. Even more noteworthy, perhaps, is the fact that in this stretch of text, S<sub>ø<sub>i</sub></sub> is separated from its antecedent by another instance of S<sub>ø</sub>, which refers to another male character, 'Lilla', who is also referred to as *he*. This intermingling between various instances of S<sub>ø</sub> and *he* referring to different entities ensure that it is no easy task to distinguish between antecedents in cases such as this. The author certainly makes no attempt to distinguish clearly between the two. It is worth noting that Miller's edition of *Bede* replaces the final empty *he* with a full NP to distinguish between antecedents in the idiomatic PDE translation, where the final line is rendered as "[a]nd Eomær thrust through the king's attendant and wounded the king" (Bede, translation 123.20). It seems excessively difficult to successfully distinguish between these antecedents when relying exclusively on masculine personal pronouns – whether overt or not. A similar example is given in (4.62), while example (4.63) illustrates an instance of S<sub>ø</sub> surrounded by two possible masculine antecedents. In these cases, the verbal morphology is of no help in identifying the antecedent. It seems substantial amounts of inference and textual interpretation on the part of the reader is needed to correctly identify the reference of S<sub>ø</sub> – and the fact is that it would probably be nearly as difficult to decipher these clauses even if the subject pronouns had been overt.

(4.62) *Lareow, ic brohte minne sunu dumbne gast hæbbende. Se swa hwær swa he hine*  
 master, I brought my son<sub>i</sub> dumb spirit<sub>j</sub> having. He<sub>j</sub> wherever he<sub>j</sub> him<sub>i</sub>  
*gelæcð [S<sub>ø</sub>]<sub>j</sub> forgnit hine, & [S<sub>ø</sub>]<sub>i</sub> toðum gristbitað & forscrinçþ;*  
 takes [he]<sub>j</sub> throws-down him<sub>i</sub>, and [he]<sub>i</sub> teeth gnashes and withers.  
 ‘Master, I brought my son, who has as dumb spirit. Wherever it takes him, it throws  
 him down, and he gnashes his teeth and withers away.’  
 (WSCp Mk. 9.18)

(4.63) *Se læce, ðonne he cymð ðone untruman to sniðanne, æresð he sceawað ðæt cumbl,*  
 the surgeon<sub>i</sub>, when he<sub>i</sub> comes the patient<sub>j</sub>-A to cut, first he<sub>i</sub> examines the swelling  
*& siððan hine tweonað ymb ðæs untruman geðyld, hwæðer he geðafian mæge ðæt*  
 and after him<sub>i</sub>-A doubts concerning the patient<sub>j</sub>’s patience, whether he<sub>j</sub> allow may that  
*hine mon sniðe.[S<sub>ø</sub>] Hyt ðonne his læceseax under his claðum oððæt he hine wundað:*  
 him<sub>j</sub> one cuts. [he]<sub>i</sub> hides then his surgeon-knife under his clothes until he<sub>i</sub> him<sub>j</sub> wounds  
 ‘The surgeon, when he comes to operate on the patient, he first examines the swelling,  
 and afterwards he has doubts concerning the patience of the patient, whether he will  
 allow that one should cut him. He then hides his scalpel under his clothes until he cuts  
 him.’  
 (CP 187.9)

Now, in cases such as the above, many medieval authors and scribes seem to have relied more heavily on use of pronouns than is common today, and it may even be said that such use of identical pronouns could be a genre feature for many medieval texts. The introduction of full NPs to break up sequences of identical referential pronouns would ease the process of identification, as demonstrated by the insertion of ‘Eomær’ in Miller’s translation of *Bede*. Ease of interpretation seems not to have been an issue for the author, but it could be that this reliance on pronouns could have been motivated by spatial restraints, as writing materials such as pens, ink and vellum were luxury goods. However that may be, the examples above serve to illustrate the difficulties involved in correctly identifying the antecedent. It is hard to see any syntactic systematicity at work here, and semantic inference and pragmatic knowledge seems to be the main method of deciphering instances of S<sub>ø</sub> similar to those seen in (4.61)–(4.63).

Another possibility is that the prominence of the figures appearing in the text may have served some function in determining whether subject pronouns could be left empty. That is, it could be speculated that only pronouns referring to the most prominent characters could be left empty. We may then hypothesise that the identity of these empty subjects is recoverable through representing “who the narrative is about”. In example (4.34), for instance, it was seen that S<sub>ø</sub> in fact did not correspond with the immediately preceding third person NP, as would perhaps be expected if syntactic “redundancy” or ease of identification was the main motivation for utilising an empty pronoun, but rather with an NP occurring at

some distance. The antecedent in question, ‘Augustine’, can be argued to be the focus of the discourse, as the entire section of text revolves around his journey to England. It may be that it is the status as theme or topic that allows a referential pronoun referring to ‘Augustine’ to be deleted here. Consequently, then, the reader of this section in *Bede* would realise – due to his pragmatic knowledge – that ‘Augustine’ is the referent of the empty pronoun, on basis of being the focus of the discourse. If this concept plays a role in identifying the antecedent, it would be expected that instances of  $S_\emptyset$  are regularly co-referent with the most prominent figure in the discourse.

The notion that prominent discourse figures could serve as an aid in identifying the antecedent ties into the concepts of *theme* and *topic*. These concepts have traditionally been somewhat interchangeably used to denote what the sentence or discourse “as a whole is about” (Matthews 2007: 407, 410–411). However, a distinction should be made between *topic* or *theme* in the pragmatic sense and *topic* in the syntactic sense primarily utilised by generative scholars – namely to indicate the entity occupying the clause-initial position. As mentioned in Chapter 2, *topic-drop* has been suggested as an explanation for what appears to be “subjectless” structures in languages such as Old Norse and Vietnamese.

However, the citations featuring  $S_\emptyset$  analysed in this work show quite clearly that this hypothesis is erroneous – or at least not systematic, and thus inadequate at best. While the examples discussed above illustrate examples of empty pronouns referring to the most prominent figures, there are plenty of examples which do not conform to this pattern, or is in direct opposition to it. As a case in point, take example (4.61). Here, there are two instances of  $S_\emptyset$ , where one refers to ‘Eomer’ and the other to ‘Lilla’. Deciding which of them is the most prominent figure is virtually impossible, as they both play central parts in the discourse. Regardless, that point is moot, as ‘Eomer’ and ‘Lilla’ are *both* represented by empty pronouns. The same is the case in example (4.62), where there are also two instances of  $S_\emptyset$ , referring to the ‘dumb spirit’ and the ‘son’, respectively. It is impossible to decide which of them is the more prominent figure, and even if there were a way of doing so, it is inconsequential, as both of them are left empty.

Allen (1999) touches upon something very similar to what has been discussed above when dismissing what she calls the ‘thematicity hypothesis’. She lists as a possible explanation for *coordinate subject deletion* that what determined the permissibility of utilising an empty subject pronoun may not have been this pronoun’s co-referentiality with another subject, but rather the empty pronoun’s status as “theme of discourse”. However, she rejects this explanation on the grounds that it is difficult to “define exactly what is meant by ‘theme

of discourse' in such a way as to account for the numerous counterexamples, and also explain why "the controller of CSD [i.e. the antecedent - KAR] is usually a subject" (Allen 1999: 55–56). She concludes her discussion by saying that while thematicity may not have been the determining factor allowing the non-expression of a subject pronoun, pragmatics may have interacted with grammatical features to a greater extent than in PDE. Thus, while dismissing the "thematicity hypothesis", she states that it does appear "that discourse factors could sometimes interfere with what had become a grammatically controlled process" (Allen 1999: 58). This is an appealing notion, and it constitutes a fine point to leave off the current discussion. It seems fitting to conclude that a character's prominence or importance as theme or topic of discourse may have had some influence on  $S_\emptyset$ , in combination with grammatical factors. It is possible that such thematicity could be of some aid in identifying its antecedent, in certain cases – yet the counterexamples are numerous and it would certainly be wrong to imply that this is a main factor in the "subjectless" clauses of OE.

As a final note, however, one should also consider the inconvenient circumstance that, due to the fact that many of these texts are translations, influences from Latin syntax may be responsible for some of the occurrences of  $S_\emptyset$  discussed here. The texts which feature the highest text-internal frequencies of  $S_\emptyset$  compared to  $S_{\text{pron}}$ , *Bede* and *Orosius*, are both translations of Latin originals. As was said in Chapter 3, though, it is difficult to determine whether instances of  $S_\emptyset$  are present because of the Latin original, or whether they were utilised by OE scribes who felt the constructions were acceptable in the vernacular as well. As referenced in Chapters 2 and 3, evidence collected from Gothic and OHG suggests tentatively that scribes would insert an overt pronoun if the empty variant was judged to be incompatible with vernacular idiom. The only way to rigorously check this would be to compare every translated OE text with its Latin counterpart. In addition, it would in principle be possible to compare frequencies for  $S_\emptyset$  in translations and non-translations – yet there are too few of the latter in OE. In either case, this lies beyond the scope of this work, and has consequently not been done. Thus, the frequencies detailed in this study will by necessity have to be taken "at face value", but one should remain conscious of the possibility of influence from Latin.

## **4.10 Closing discussion**

### **4.10.1 Predominance in second conjunct clauses**

As demonstrated above, the majority of the instances of  $S_\emptyset$  analysed here occur in second conjunct clauses. While that in itself is unremarkable, and may indicate simply that there is an



overweight of second conjunct clauses in the selected corpus, it was demonstrated by comparing empty and overt subject pronouns according to clause type across the corpus texts, that  $S_{\emptyset}$  in fact occurs more frequently in second conjunct clauses than in any other clausal environment. As remarked above, this corresponds quite well with the observed distribution of  $S_{\emptyset}$  in languages such as Gothic, OHG and ON, as these languages primarily utilised empty subject pronouns in main clauses, whether of the second or non-conjunct variety. It was also observed that for OE, second conjuncts is the only clause type in which  $S_{\emptyset}$  could at all be claimed to be an “active” phenomenon in the language – and barely so, at an overall relative frequency of 2.4%. Additionally, the relatively high frequencies for  $S_{\emptyset}$  in this clause type are not restricted only to *Bede*, but are rather quite evenly distributed among the corpus texts. In all other clause types, the overall frequency for  $S_{\emptyset}$  was observed to be less than 1%, and in most of the clause types according to texts, the frequency was actually closer to 0.5%. This means that if a “grammatical competition” between  $S_{\emptyset}$  and  $S_{\text{pron}}$  similar to the one Axel (2005: 28) assumes for OHG also took place in the English language, the contest was completed by the period of OE – and overt subject pronouns left as the “winner” (section 2.8.2). Indeed, these low frequencies lead to questioning whether some scholars may not have overestimated the occurrence rates of “subjectless” clauses in OE. The data collected here certainly does not agree with for instance van Gelderen’s claim that “[i]n Old English, pro-drop is quite common” (van Gelderen 2000: 121).<sup>13</sup> It is also impossible to agree with Pogatscher (1901) that “the pronoun in OE subordinate clauses need not be expressed”. It was suggested above that the reason this linguistic phenomenon may have been exaggerated by some previous scholars is the lack of data in this area of research into OE.

As for explaining the occurrence of the phenomenon, given the observed dominance in second conjunct clauses, it seems to be no stretch of the imagination that one of the deciding factors involved in the majority of the clauses featuring  $S_{\emptyset}$  discussed in this work is related to the rules of OE governing the situational permissibility of deletion of a clause element under coordination. Contrary to the modern language, it seems that in some rare cases, subject omission occurs in second conjunct clauses regardless of whether the empty pronoun is co-referent with a preceding subject and regardless of whether the antecedent actually occurred in the immediately preceding main clause.

---

<sup>13</sup> Indeed, van Gelderen (2000) claims to have “illustrated that pro-drop occurs regularly in Old English”, that “pro-drop is quite common” and that this is indicative of “the strength of the verbal person features” (van Gelderen 2000: 121, 149). Based on the figures discussed in section 4.7.1, I do not see how that could possibly be the case, as the frequencies for  $S_{\emptyset}$  are very low. Also, as was seen in section 4.9, it is highly doubtful whether the OE verbal inflections are as detailed as van Gelderen claims.

The question remains, though, whether our examples truly reflect cases where the situational syntactic criteria for deletion of a subject pronoun were met, or whether some – or many – of the 267 instances of  $S_{\emptyset}$  discussed here in fact represent grammatical errors. It should be kept in mind that the total number of occurrences is very low, both in observed and relative frequencies. Also, it must be fair to say that the instances of  $S_{\emptyset}$  are quite erratically distributed, with no apparent text-internal systematicity, in the sense that parallel structures to those featuring  $S_{\emptyset}$  actually have  $S_{\text{pron}}$  in the overwhelming majority of the cases. If the rules of OE regarding subject omission had been the same as in PDE – and they do correspond in very many cases – it is all too easy to imagine a situation where a scribe, noting that his sentence starts with a coordinating conjunction, determines to leave the subject empty in the mistaken belief that the subject pronoun in the second conjunct was co-referent with that of the immediately preceding main clause. It may be that some scribes have simply “lost track” of both the clausal location and function of the antecedent. Allen (1999) mentions this possibility, stating that the “majority of the exceptional examples [...] involve examples in which it would be particularly easy for the author or scribe to forget that the subject of the coordinated clause was not in fact the grammatical subject of the first conjunct” (Allen 1999: 57).

However, while the possibility exists that some or many of our examples of  $S_{\emptyset}$  are the result of scribal errors, it cannot be ruled out that these examples actually represent cases where OE may have had less stringent rules for element deletion under coordination than the modern language. The phenomenon represented by  $S_{\emptyset}$  is certainly rare and appears very erratically, yet it seems too widespread to be entirely coincidental – especially in certain texts, such as for instance *Bede*. The fact also remains that examples of  $S_{\emptyset}$  were found in all the corpus texts used in this work. It would then seem unlikely that all or many of the instances of  $S_{\emptyset}$  treated in this work are consequences of slips on the part of the author or scribe. Thus, the position adopted here is that the high frequency of  $S_{\emptyset}$  in second conjuncts is indicative of some degree of systematicity, or remnant of systematicity, in that these examples may reflect an older set of rules for subject omission in early Germanic which by the time of the extant texts had fallen into disuse in OE. One should recall here that the extant Gothic manuscripts and some Old High German texts feature very high frequencies for  $S_{\emptyset}$  (cf. section 2.7). Unfortunately, however, it seems very difficult, if not impossible, to achieve absolute certainty here.

#### 4.10.2 Subject-antecedent predominance

As section 4.7.2 indicates,  $S_\emptyset$  is by far more frequently co-referent with a subject occurring in the surrounding context, or alternatively, part of a subject, or a subject in combination with an additional element, such as an object or object of a preposition. The co-referent subject is located in a preceding clause in the overwhelming majority of the cases, but in three of my citations, the antecedent is located in a following main or subordinate clause. It was seen that between 56.7% and 66.0% of the instances of  $S_\emptyset$  discussed here are co-referent with a single preceding or following nominal or pronominal subject.<sup>14</sup> The frequencies for  $S_\emptyset$  co-referent with a subject occurring in the surrounding context may seem slightly low, but it should be remembered that “regular” deletion of co-referent subjects occurring in second conjunct clauses, as it is still practised in PDE, is outside the scope of this thesis (cf. section 3.3). On this basis, it really must be said that the frequencies ranging from 56.7%–66.0% are actually quite high, and there can be no doubt that this is the dominant antecedent type. Combined with the tendency of  $S_\emptyset$  to occur more frequently in second conjuncts, as discussed in the preceding section, the fact that the antecedents of the empty subjects discussed here are co-referent with a subject in the majority of the cases could constitute additional evidence for the idea that OE at some stage may have had slightly different rules regarding the permissibility of subject omission in coordinate clauses. As stated above, the instances of  $S_\emptyset$  are so few in my corpus of texts that it seems doubtful that these rules should still be “active” in the language at the OE stage, yet as also mentioned above, it is possible that the instances we do see of this phenomenon may reflect remnants or relics of an older grammatical system governing subject omission under coordination.

#### 4.10.3 Distance/identification

The examples given in section 4.9 should demonstrate that identification of the antecedent is not necessarily a straightforward matter. First of all, data collected here shows that  $S_\emptyset$  is not restricted to occur in close proximity to its antecedent, as the empty subjects which had an overt antecedent were quite evenly distributed over the four identified categories indicating the distance separating  $S_\emptyset$  and its antecedent. If it is the case that the antecedent should be easily identifiable in order for subject omission to take place, one would assume that the empty pronoun would occur in very close proximity to its co-referent expression. As we have seen, this is not the case in the corpus examined here, and identifying the antecedent is in

---

<sup>14</sup> These figures pertain to the instances of  $S_\emptyset$  discussed in section 4.7 and onwards. This means that instances of  $S_\emptyset$  occurring in hortative subjunctive constructions and empty subject relatives are not included here.

many cases quite difficult. It was seen that the verbal morphology could be of some help, in that it quite efficiently distinguishes between the singular and the plural, and to a certain extent between person, but that there is too much syncretism for the system to systematically disambiguate the antecedent. This means, then, that the “identification hypothesis” has not been corroborated. It also seems that many of the problematic cases in terms of identification involve several referents of the same person and number, namely third person singular referents, and thus even a system with unambiguous morphological marking of the verb would be of little help. In fact, it seems clear that in many cases, the clause would be equally confounding if an overt subject pronoun had been present. However that may be, it is certainly not the case here, as the ideal state would be for the generativists investigating *pro*-drop, that verbal inflections make an overt subject pronoun recoverable and redundant. In addition, it was also observed that there is often confusing alternation between referents, which can greatly complicate the process of identification. Examples were also given where instances of  $S_0$  occurring in close proximity to each other refer to separate entities, with no way to distinguish grammatically between them. In cases such as (4.61)–(4.63), it seems the only way to identify the correct antecedents is through pragmatic considerations.

It was further suggested that the prominence of the characters appearing in the texts could have some bearing on the identification of the antecedent, in that such characters could function as the theme or topic of discourse. If this was a factor involved in subject omission, one would expect instances of empty pronominal subjects to correspond with any such prominent figures. This would in turn mean that whenever an empty pronoun is encountered, the reader might be able to assume that the “gap” refers to the most prominent figure in the discourse through this entity’s capacity of theme or topic, and thus have no problem identifying the antecedent. However, clear counterexamples were given to this “prominent figures” hypothesis, and it is clear that  $S_0$  does not necessarily refer to the figures in the texts representing the discourse topic/theme. It was concluded that thematicity could very well play a role in some instances of  $S_0$ , and thus be of some aid in identifying the antecedent, but that this is probably not a major influence on the permissibility of empty subject pronouns. However, it seems likely that pragmatics and “text-reading skills” must have been one of the key factors in deciphering clauses featuring these empty subjects. When thematicity is of little help, and verbal inflections help to some extent only, this seems to be the final method of correctly identifying the antecedent – and thus also the meaning of the clause. Finally, the possibility of influences from Latin was mentioned.

#### 4.11 Summary

This chapter has presented and discussed the data collected for this thesis, thus forming the main part of this study. First, a presentation was given of the relative distribution of  $S_\emptyset$  in the corpus of texts, compared and contrasted with the occurrences of  $S_{\text{pron}}$  in the same texts. It was seen that the frequencies for  $S_\emptyset$  were very low. In the comparison between  $S_\emptyset$  and  $S_{\text{pron}}$ , the frequencies ranged between 0.3%  $S_\emptyset$  (ByrM) and 3.3%  $S_\emptyset$  (*Bede*). The sole exception was Alfred's *Laws* which, due to its reliance on hortative subjunctive structures, featured a frequency of 39.1%  $S_\emptyset$ . The difference between eOE and IOE was statistically significant, even when LawAf was discounted. This means that it is possible to discern a development toward the permanent loss of empty pronominal subjects in the two periods of Old English. However, it is unlikely that the figures illustrate a major shift in OE from being a language that permitted empty subject pronouns to one that did not, as the instances of  $S_\emptyset$  are very few in both periods. Further, it was seen that the analysed tokens of  $S_\emptyset$  fall into three main categories:  $S_\emptyset$  in hortative subjunctive structures, empty subject relatives and one final, less uniform category. The latter is by far the largest and complex category, and the main part of this chapter was dedicated to describing these 267 instances of  $S_\emptyset$ . This description was presented as a quantitative analysis of the role of a number of structural variables proposed by previous research. These included the types of clause in which  $S_\emptyset$  occurs, the function of its antecedent, the location of the antecedent and the distance between  $S_\emptyset$  and its antecedent. By comparing the clausal distribution of both  $S_\emptyset$  and  $S_{\text{pron}}$  in a variety of clause types, it was seen that  $S_\emptyset$  most frequently occurs in second conjunct clauses. The data presented led to concluding that some earlier scholars of OE may have been too quick in labelling “subjectless” clauses a “native” feature of OE, as the occurrences are so very few. It was confirmed that  $S_\emptyset$  is generally co-referent with a subject occurring in the surrounding context, or with a subject in combination with another element. It was also observed, however, that  $S_\emptyset$  could be co-referent with many types of antecedent, including direct and indirect objects, objects of prepositions, in addition to only parts of elements. A section was then dedicated to the difficulties involved in identifying the antecedent, and thereby understanding the meaning of the clause. It was concluded that there is little uniformity to be seen among the instances of  $S_\emptyset$  discussed in the main part of this chapter, and also that it is very difficult to determine any systematicity in the examples discussed here. It seems impossible to draw conclusions based on what “allowed” or governed  $S_\emptyset$  in OE, but it has been suggested here that the extant occurrences may be the result of a combination of factors. First, it may be assumed that an antiquated early Germanic grammatical system permitting empty subject pronouns can be

held accountable for some of the instances; secondly, pragmatic factors may have factored in, and finally; it may not be ruled out that some of the instances of  $S_\phi$  seen here may be the result of scribal error.

## CHAPTER 5

### SUMMARY AND CONCLUSION

The present thesis has been concerned with the non-expression of referential subject pronouns and the resulting “subjectless” sentences occurring on Old English prose. The study has investigated 450 occurrences of such sentences taken from a corpus consisting of eleven Old English texts. The aim of the investigation was to describe the occurrence of empty referential subject pronouns, and determine whether any systematicity could be seen with regard to their syntactic behaviour. While the study has been focussed primarily on syntactic aspects of these so-called empty subjects, certain pragmatic factors have also been taken into consideration.

Chapters 1, 2 and 3 provided the background for the study. Chapter 1 accentuated the distinction between languages which require overt referential pronominal subjects in the clause structure and languages which do not. It was pointed out that the modern Germanic languages, including English, require pronominal subjects to be overt in all but a few syntactic environments. However, it was also noted that all the Germanic languages are assumed to have developed from a stage where such pronominal subjects could be realised as empty, in a fashion comparable to genuine *pro*-drop languages such as Spanish and Italian. In this connection, it was seen that some scholars, notably Baker (2007), Mitchell (1985), Traugott (1992) and van Gelderen (2000), assume that empty subjects were a more or less idiomatic feature of Old English, as well as in Gothic and Old High German. Still, it was noted that little consensus as to the extent of this phenomenon has been reached, as evidenced by the widely differing accounts of van Gelderen (2000: 149) – who claims that “*pro*-drop occurs regularly in Old English” and Mitchell (1985: 633) – who says that the idiom appears “only spasmodically”. On background of this, it was stated that the main purpose of this work would be to quantify this understudied phenomenon.

It was decided that the focus of the study was to be explorative, and primarily aimed at providing a quantitative analysis of structural variables proposed by previous scholarship as being relevant to subject omission in OE. This decision was made on background of the lack of solid data regarding the extent of the phenomenon in OE. Two research questions aimed at facilitating study of empty subjects in Old English were formulated, the phrasing of which indicating that certain hypotheses would be tested implicitly, including, for instance, whether empty pronominal subjects were as frequent as suggested by e.g. Baker (2007), Mitchell (1985) and van Gelderen (2000). Additionally, it would be tested whether whether the

syntactic environments in which the empty subject occurred had any bearing on the permissibility of this phenomenon in OE.

Chapter 2 presented previous research on empty subjects. First, an account was given of the generative investigations of the phenomenon under the labels of *pro-drop* or *null subject*. A generative study of empty subjects in Old Icelandic was offered before two of the most central studies of empty subjects in Old English were presented. Short overviews were also provided of the status of research into empty subjects in Gothic and Old High German. Chapter 3 was concerned with method and data collection, and also discussed various problems of analysis.

Chapter 4 comprised the main part of this study, where the collected data was analysed and discussed. First, it was observed that the frequencies for empty subjects in Old English prose were very low, but also that the distribution of the phenomenon was significantly greater in texts belonging to the early Old English period.

It was also observed that the 450 instances of empty subjects could be separated into three main groups. Two of these groups were easily classified. These included cases where hortative subjunctive verbs combined with empty subjects and cases where subject relative pronouns were left empty. The final category was much less uniform. The majority of Chapter 4 was dedicated to classifying these instances of empty subjects according to various syntactic and pragmatic criteria. These criteria were first referenced in conjunction with the research questions in Chapter 1, and detailed in full in Chapter 3. It was seen that empty subjects were most frequent in second conjunct clauses, yet even here the frequencies were low on a text-to-text basis. While it was noted that the empty pronoun could be co-referent with many types of antecedent, it most frequently shared reference with a subject occurring in the preceding context.

The overall low frequencies led to concluding that some scholars, van Gelderen in particular, may have exaggerated the distribution of empty pronominal subjects in Old English. At several stages it was commented that the frequencies were so low that the phenomenon of subject omission must be characterised as nearly dead by the extant Old English period.

In terms of why the phenomenon occurs in Old English, no satisfactory conclusion was reached. It was shown that the the notion that the identity of referent should be clear from verbal inflections was not systematically valid. In some cases, verbal inflections were admittedly helpful – particularly in differentiating between the singular and the plural – yet in many cases, the syncretism of the verbal system leads to the combination of empty subjects



with ambiguously inflected verbs – which, ideally, should be impossible if it really is true that the verb makes a pronominal subject recoverable, and thus redundant. In this context, even the plural may be considered ambiguous, as most verbs inflected for the plural have an *-on* or *-an* ending. This means that even though it is clear that some plural entity constitutes the antecedent, the identity of the referent may still not be entirely clear.

It was also seen that the main problem in identifying the antecedent did not necessarily depend on having available unambiguous verbal inflections. In a considerable number of cases, it was seen that several referents of the same number and person intermingled with no structural means of differentiating between them. In these cases, the clause would actually be equally inexplicable even with the presense of an overt pronoun. In such cases, it seems only that pragmatic factors may reveal the referent of the empty subject.

Based on the data presented in this thesis, then, it was concluded that it was impossible to reach an overarching conclusion regarding what “governed” empty referential pronominal subjects in Old English. It was, however, speculated that the tokens discussed here could be explained by a combination of factors, of which one was pragmatic features. It was also suggested that some may represent genuine remnants of an antiquated Germanic grammatical system, and that some instances may simply be explainable by scribal or authorial error.

While empty pronominal subjects have been studied extensively by generativists investigating the *pro*-drop parameter, this thesis is to the best of my knowledge the first in-depth study of empty subjects in Old English. Its main merit is that it concretises a formerly understudied and somewhat impressionistic area of English historical syntax. The presented results also show that such a study was needed, particularly in order to dispel some misconceptions about the extent of the phenomenon in the period *c.*700–*c.*1100. While I do not believe that the last word has been said regarding empty referential pronominal subjects in Old English, it is hoped that this work forms a basis on which future studies can be built.

## REFERENCES

### Primary sources

- Two of the Saxon Chronicles Parallel*, ed. by C. Plummer. 1965. Oxford: Clarendon Press. Reissued D. Whitelock, Oxford 1952.
- The Old English Version of "Bede's Ecclesiastical History of the English People"*, ed. by T. Miller. 1959–1963 (1890–1898). EETS 95, 96, 110, 111. London: OUP.
- King Alfred's Old English Version of Boethius de Consolatione Philosophiae*, ed. by W. J. Sedgefield. 1899. Oxford: Clarendon Press. Reprinted Darmstadt 1968.
- King Alfred's West-Saxon Version of Gregory's Pastoral Care*, ed. by H. Sweet. 1958 (1871). EETS 45, 50. London: OUP.
- The Old English Orosius*, ed. by J. Bately. 1980. EETS s.s. 6. London: OUP.
- Lieberman, Felix. 1903–1916. *Die Gesetze der Angelsachsen*. Halle. Reprinted Aalen 1960.
- Ælfric's Catholic Homilies. The First Series*, ed. by P. Clemoes. 1997. EETS s.s. 17. Oxford: OUP.
- Ælfric's Catholic Homilies. The Second Series*, ed. by M. Godden. 1979. EETS s.s. 5. London: OUP.
- Ælfric's Lives of Saints*. Ed. by Walter William Skeat. 1966 (1881-1900). EETS 76, 82, 94, 114. London: OUP.
- The Four Gospels in Anglo-Saxon, Northumbrian and Old Mercian Versions*, ed. by W. W. Skeat. 1871–1887. Cambridge: CUP. Reprinted Darmstadt 1970.
- Bischof Warferth von Worcester Übersetzung der Dialoge Gregors des Grossen*, ed. by H. Hecht. 1965 (1900–1907). Bibliothek der Angelsächsischen Prosa, V. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Byrhtferth's Enchiridion*, ed. by P. S. Baker and M. Lapidge. 1995. EETS s.s.15. Oxford: OUP.

### Secondary sources

- Abraham, Werner. 1993. "Null Subjects in the History of German: From IP to CP". *Lingua* 89, 117–142.
- Allen, Cynthia L. 1999. *Case Marking and Reanalysis. Grammatical Relations from Old to Early Modern English*. Oxford: OUP.
- Axel, Katrin. 2005. "Null Subjects and Verb Placement in Old High German". In *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives*, eds. Stefan Kepser and Marga Reis. New York: Mouton de Gruyter.

- Baker, Peter S. 2007. *Introduction to Old English*. Oxford: Blackwell.
- Bately, Janet. 1991. "The Nature of Old English Prose". In *The Cambridge Companion to Old English Literature*, eds. Malcolm Godden and Michael Lapidge. Cambridge: CUP.
- Barðdal, Jóhanna. 2000. "Oblique Subjects in Old Scandinavian". *North-Western European Language Evolution* 37, 25–51.
- Bergen, Linda van. 2000. "The Indefinite Pronoun *man*: 'nominal or 'pronominal'?" In *Generative Theory and Corpus Studies. A Dialogue from 10 ICEHL*. Topics in English Linguistics 31, 103–22. Berlin/New York: Mouton de Gruyter.
- Chafe, Wallace. 1976. "Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View". In *Subject and Topic*, ed. Charles N. Li. New York: Academic Press.
- Clement, Richard W. 1986. The Production of the Pastoral Care: King Alfred and his Helpers." In *Studies in Earlier English Prose*, ed. Paul E. Szarmach. New York: State University of New York Press.
- Eggenberger, Jakob. 1961. *Das Subjektspronomen in Althochdeutschen*. Doctoral diss., University of Zürich.
- Fertig, David. 2000. "Null Subjects in Gothic". *American Journal of Germanic Linguistics & Literatures* 12(1), 3–21.
- Fulk, R. D. and Christopher M. Cain. 2003: *A History of Old English Literature*. Oxford: Blackwell.
- Garmonsway, G. N. 1954. *The Anglo-Saxon Chronicle*. London: Dent.
- Gelderen, Elly van. 2000. *A History of English Reflexive Pronouns. Person, Self and Interpretability*. Amsterdam: John Benjamins.
- Giles, J. A. 1958 [1969]. "The Laws of King Alfred the Great". In *The Works of King Alfred the Great*, eds. J. A. Giles *et al.* Vol. III, 119–40. New York: AMS Press.
- Gilligan, Gary M. 1987. *A Cross-Linguistic Approach to the Pro-Drop Parameter*. PhD diss., University of Southern California.
- Gundel, Jeanette K. 1980. "Zero NP-Anaphora in Russian: A Case of Topic-Prominence". In *Papers from the Parasession on Pronouns and Anaphora*, eds. Jody Kreiman and Almerindo E. Ojeda. Chicago: Chicago Linguistic Society.
- Haegeman, Liliane. 1996. *Introduction to Government and Binding Theory*. 2<sup>nd</sup> ed. Oxford: Blackwell.
- Haugland, Kari E. 2007. *Old English Impersonal Constructions and the Use and Non-use of Nonreferential Pronouns*. Doctoral diss., University of Bergen.

- Heggelund, Øystein I. 2010. *Word Order in Old English and Middle English Subordinate Clauses*. Doctoral diss., University of Bergen.
- Held, Karl. 1903. *Das Verbum ohne Pronominales Subjekt in der älteren Deutschen Sprache*. Palaestra 31. Berlin: Mayer & Müller. (Repr. 1967, New York: Johnson Reprint Corp.)
- Huang, Yan. 2007. *Anaphora. A Cross-linguistic Study*. Oxford: OUP.
- Jaeggli, Osvaldo and Kenneth J. Safir (eds.). 1989. *The Null Subject Parameter*. Dordrecht: Kluwer Academic Publishers.
- Kemenade, Ans van. 1987. *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris
- Kohonen, Viljo. 1978. *On the Development of English Word Order in Religious Prose around 1000 and 1200 A.D. A Quantitative Study of Word Order in Context*. Doctoral diss., Åbo: Publications of the Research Institute of the Åbo Akademi Foundation.
- Koopman, Willem F. 1997. "Another Look at Clitics in Old English". *Transactions of the Philological Society* Volume 95:1, 73–93.
- Kraus, Carl. 1894. *Deutsche Gedichte des Zwölften Jahrhunderts*. Halle: Max Niemeyer.
- Leech, G. 1992. "Corpora and Theories of Linguistic Performance". In *Directions in Corpus Linguistics*, ed. J. Svartvik. Berlin: Mouton de Gruyter.
- Liuzza, R. M. 2000. *The Old English Version of the Gospels. Vol 2: Notes and Glossary*. EETS: Oxford: OUP.
- Magoun, Francis Peabody, jr. 1935. "Territorial, Place-, and River-Names in the Old English Chronicle, A-Text". *Harvard Studies and Notes in Philology and Literature*, 69–111.
- Matthews, P.H. 2007. *The Concise Oxford Dictionary of Linguistics*. Oxford: OUP.
- McEnery, Tony et al. 2006. *Corpus-Based Language Studies. An Advanced Resource Book*. New York: Routledge.
- Mitchell, Bruce. 1985. *Old English Syntax*. Vols. 1–2. Oxford: Clarendon.
- Ohlander, U. 1943–1944. "Omission of the Object in English", *Studia Neophilologica* 16, 105–127.
- Paul, Hermann. 1919. *Deutsche Grammatik*. Vol. 3, part 4. Halle: Max Niemeyer.
- Pogatscher, Alois. 1901. "Unausgedrücktes Subjekt im Altenglischen". *Anglia* 23, 261–301.
- Pulsiano P. and Elaine Treharne (eds.). 2001. *A Companion to Anglo-Saxon literature*. Oxford: Blackwell.
- Quirk, Randolph and C. L. Wrenn. 1976. *An Old English Grammar*. London: Methuen.

- Robinson, Orrin W. *Old English and its Closest Relatives. A Survey of the Earliest Germanic Languages*. London: Routledge.
- Rosén, Victoria. 1998. *Topics and Empty Pronouns in Vietnamese*. Doctoral diss., University of Bergen.
- Sigurðsson, Halldór Á. 1993. “Argument-drop in Old Icelandic”. *Lingua* 89, 247–280.
- Sonderegger, Stefan. 1974. *Althochdeutsche Sprache und Literatur: Eine Einführung in das Älteste Deutsch*. Berlin: de Gruyter.
- Streitberg, Wilhelm. 1906. *Gotisches Elementarbuch*. Heidelberg: Carl Winter (2<sup>nd</sup> ed., 1<sup>st</sup> ed. 1897).
- Thorpe, Benjamin. 1843. *The Homilies of the Anglo-Saxon Church. 1*. London: Richard and John E. Taylor.
- Traugott, Elizabeth Closs. 1992. “Syntax”. In *The Cambridge History of the English Language*, Vol. 1: *The Beginnings to 1066*, ed. Richard M. Hogg, 167–286. Cambridge: Cambridge University Press.
- Visser, F. Th. 1963–73. *An Historical Syntax of the English Language*. Vol. 1. Leiden: E. J. Brill.
- Wrenn, C. L. 1967. *A Study of Old English Literature*. London: George Harrap.
- Zimmermann, Michael. 2009. “On the evolution of expletive subject pronouns in Old French”. In *Proceedings of the Workshop “Null subjects, expletives and locatives in Romance”*, eds. G. A. Kaiser and E.–M. Remberger. Arbeitspapier 123. Fachbereich Sprachwissenschaft, Universität Konstanz.

### Internet sources

<http://gandalf.uib.no/lingkurs/webroot/index.php?page=corpus/collections&lang=en&course=lingmet> (collected 23 July 2010)

<http://tapor.library.utoronto.ca/doecorpus/> (Dictionary of Old English web corpus)

<http://www-users.york.ac.uk/~lang22/YCOE/YcoeHome.htm> (YCOE homepage)

<http://www-users.york.ac.uk/~lang22/YCOE/doc/corpussearch/CSRefToc.htm> (YCOE search manual)

<http://www-users.york.ac.uk/~lang22/YCOE/doc/annotation/YcoeRefToc.htm> (YCOE manual for syntactic annotation)

<http://faculty.vassar.edu/lowry/newcs.html?> (Chi-square contingency table test)

<http://faculty.vassar.edu/lowry/csfit.html> (Chi-square goodness of fit test)

## SUMMARY IN NORWEGIAN

Denne masteroppgaven har befattet seg med ikke-uttrykte – såkalt ”tomme” – referensielle subjektspronomen i gammelengelsk prosa. Studien har undersøkt 450 forekomster av ”subjektløse” setninger tatt fra et korpus bestående av elleve gammelengelske prosatekster. Målet for undersøkelsen har vært å beskrive forekomsten av disse tomme subjektene, og avgjøre hvorvidt man kan observere noen grad av systematisitet i forhold til disses syntaktiske egenskaper. Enkelte pragmatiske faktorer har også blitt behandlet.

Studiens hovedfokus har vært kartleggende. I og med at tidligere undersøkelser i stor grad har basert seg på et meget begrenset datamateriale, har man her forsøkt å kvantifisere fenomenet i så stor grad som mulig. Den kvantitative analysen har i hovedsak basert seg på diverse strukturelle variabler som har blitt foreslått i tidligere forskning. Dermed har man også testet en rekke påstander vedrørende dette fenomenets utbredelse i gammelengelsk, samt undersøkt hvorvidt de syntaktiske omgivelsene de tomme subjektene forekommer i har hatt noen betydning.

Studien er bygget opp på følgende måte: kapittel 1 presenterer bakgrunnen for prosjektet, samt forskningsspørsmålene som danner grunnlaget for studien. Kapittel 2 presenterer tidligere forskning på området. Her inngår en presentasjon av relevant generativ forskning på det såkalte *pro-drop* parameteret, samt redegjørelser både for tomme subjektspronomen i gammelengelsk spesifikt, men også i andre gammelgermanske språk. Kapittel 3 gjør rede for metoden benyttet i forbindelse med datainnsamling og analyse. Enkelte analytiske problemer er også diskutert her. Kapittel 4 utgjør hoveddelen av studien, der datamaterialet blir presentert og diskutert. Kapittel 5 oppsummerer de foregående kapitlene og konkluderer i forhold til forskningsspørsmålene det har vært operert med.

Studien konkluderer med at tomme subjekter i gammelengelsk er et mye mer begrenset fenomen enn tidligere forskning har antydnet. Forekomstene er faktisk så få at man slår fast at fenomenet er mer eller mindre ”dødt” i gammelengelsk. Det slås fast at fenomenet er noe mer utbredt i tidlig gammelengelsk i forhold til senere gammelengelsk. Når det gjelder syntaktiske omgivelser, ble det sett at tomme subjekter i all hovedsak opptrer i såkalte ”second conjunct clauses”, altså det andre medlemmet av to hovedsetninger som er bundet sammen av en koordinerende konjunksjon. Det ble også sett at de tomme subjektene i flertallet av tilfellene er ko-referensielle med et subjekt som opptrer i den foregående konteksten. Det konkluderes også med at det er vanskelig å se noen systematikk i hva som ”tillot” eller ”styrte” disse tomme subjektene. Det ble likevel antatt at forekomstene diskutert i denne oppgaven kunne forklares ved en kombinasjon av faktorer, der pragmatikk uten tvil spiller en rolle. Det ble også foreslått at noen av forekomstene kunne representere levninger av et antikvert germansk grammatisk system. Man kunne heller ikke utelukke at noen av forekomstene rett og slett er utslag av grammatiske feil.

# APPENDIX: Example of a FileMaker record

<b>Example</b>	<input text"="" type="text" value="Beede"/>	
	<input type="text" value="88. 31"/>	<b>Reference</b>
	<input type="text"/>	<b>Manuscript</b>
<b>Gloss</b>	<input type="text" value="Among them were the first and the greatest Mellitus and Justus and Paulinus and Ruffianus. And (he) by them sent common the things all which to church worship and service necessary were...."/>	<b>Preceding context</b> <input type="text" value="Swilce Agustinus heht him eac onbeodan, þætte &lt;her&gt;-swære&gt; micel rip onweard &amp; fea worhton, ond he ða sende mid þæm foresprecan ærendwreca him. maran fullom to godcundra lare."/>
<b>Translation</b>	<input type="text" value="Augustine also bade them announce to him, that the harvest here was now plenteous, but the labourers few; so he sent him with the aforesaid envoys greater help for the teaching of the Word.} The first and chiefest among these were Mellitus and Justus and Paulinus and Ruffianus; and he sent by them all those things in general, which were"/>	<b>Context ref</b> <input type="text" value="88. 28"/>
	<input type="text" value="1-3"/> <input type="text" value="4-6"/> <input type="text" value="6-10"/> <input checked="" type="checkbox"/> <input type="text" value="10+"/>	<input type="checkbox"/> Missing relative <input type="checkbox"/> Imperative <input type="checkbox"/> Dative subject
	<input type="text" value="Antecedent (he) = Augustinus."/>	<input type="checkbox"/> 1st <input type="checkbox"/> 2nd <input checked="" type="checkbox"/> 3rd Missing pronoun person Number <input checked="" type="checkbox"/> Singular <input type="checkbox"/> Plural
	<input type="checkbox"/> Main <input checked="" type="checkbox"/> Syndetic sec conjunct <input type="checkbox"/> Asyndetic sec conjunct <input type="checkbox"/> Relative <input type="checkbox"/> Averbial <input type="checkbox"/> That-clause	<input type="checkbox"/> Following main <input type="checkbox"/> Following subordinate <input type="checkbox"/> Preceding subordinate <input type="checkbox"/> Indefinite <input type="checkbox"/> Misc
	<input checked="" type="checkbox"/> S <input type="checkbox"/> O <input type="checkbox"/> ObjP <input type="checkbox"/> SC <input type="checkbox"/> Indefinite <input type="checkbox"/> Gen <input type="checkbox"/> Misc	<b>Location of antecedent</b> <input checked="" type="checkbox"/> Preceding main <input type="checkbox"/> Following main <input type="checkbox"/> Preceding subordinate <input type="checkbox"/> Following subordinate <input type="checkbox"/> Indefinite <input type="checkbox"/> Misc
	<input type="checkbox"/> Grammatical function of antecedent	