PUBLICATIONS OF
THE UNIVERSITY OF EASTERN FINLAND

Dissertations in Education, Humanities, and Theology

IZABELA CZERNIAK

ANGLO-SCANDINAVIAN LANGUAGE CONTACTS AND WORD ORDER CHANGE IN EARLY ENGLISH

Anglo-Scandinavian Language Contacts and Word Order Change in Early English

IZABELA CZERNIAK

Anglo-Scandinavian Language Contacts and Word Order Change in Early English

Publications of the University of Eastern Finland
Dissertations in Education, Humanities, and Theology
No 85

University of Eastern Finland
Joensuu
2016

Grano Oy
Jyväskylä, 2016
Sarjan toimittaja: Maija Könönen
Myynti: Itä-Suomen yliopiston kirjasto
ISBN: 978-952-61-2114-7 (nid.)
ISSNL: 1798-5625
ISSN: 1798-5625
ISBN: 978-952-61-2115-4 (PDF)
ISSN: 1798-5633 (PDF)

Czerniak, Izabela
Anglo-Scandinavian Language Contacts and Word Order Change in Early English
Joensuu: University of Eastern Finland, 2016, 208 pages
Publications of the University of Eastern Finland
Dissertations in Education, Humanities, and Theology; 85
ISBN: 978-952-61-2114-7 (print)
ISSNL: 1798-5625
ISSN: 1798-5625
ISBN: 978-952-61-2115-4 (PDF)
ISSN: 1798-5633 (PDF)

ABSTRACT

This study investigates one of the most important changes that the English language underwent in the earlier course of its development – the emergence of a rigid SVO word order. Both internal and external factors have been identified in the literature as influential in the change. Among the latter, contacts with the early Scandinavian population have often been mentioned as providing an important early input. These contacts have also been regarded as one of the factors contributing to the erosion of case inflections, a change implicated in the gradual stabilisation of the SVO order. The main objective of the present study is to assess the role of these external pressures in the establishment of the new syntactic conditions in early English. In a more general perspective, this study evaluates the significance of language contacts in promoting changes in morphosyntax. This research also examines the relevance of an influential theoretical model used in the literature to explain the changes at issue, viz. Johannes Schmidt's wave theory.

The material for the present study consists of two parsed corpora representing Old and Middle English: *The York-Toronto-Helsinki Corpus of Old English Prose* (YCOE) and *The Penn Parsed Corpus of Middle English* (PPCME2, Second Edition). Together, these databases comprise almost 3 million words and contain texts from different dialects. The frequency of occurrence of the (S)VO word order in particular dialect sectors is measured and compared from various angles relevant to the study. They include distinct clause settings, focusing on nominal or pronominal constituents of NPs in word order sequences, the potential impact of the date and genre of texts in sets, as well as exploring the differences in word order distributions between texts that are translations from foreign originals and those that represent native material. The multifaceted analysis of data aims, among other matters, to evaluate the usefulness of parsed diachronic corpora in tracking large-scale linguistic changes.

The results show that (S)VO developed faster in the dialects of the areas affected by the contact, viz. the North and the East Midlands. This feature seems well established especially at the subordinate clause level. Furthermore, the highest normed frequency values for sequences with pronouns and nouns alike were found in the northernmost dialects. Both findings suggest a more external rather than internal motivation for the structural change. Evidence from genetic and archaeological stud-

ies, too, speaks in favour of a clearly marked Scandinavian zone, which temporally extended beyond the so-called Viking Era. With a repeated pattern of the most frequently occurring and most regularly distributed (S)VO particularly in the North, the connection between morphological simplification and the emerging new word order is more than likely. The prominence of the North within the dialectal spectrum likewise points to the existence of a focal area, which provided the starting point of linguistic innovations in a way advocated by the wave model. Set against the socio-political reality of medieval England, the results confirm the existence of a north-south divide, with the former constituting an auspicious setting for fostering changes of all kinds, including those occurring in the language.

Keywords: English language, word order, inflectional morphology, language contacts, dialects, variation, corpus study

Czerniak, Izabela
Anglo-Scandinavian Language Contacts and Word Order Change in Early English
Joensuu: University of Eastern Finland, 2016, 208 sivua
Publications of the University of Eastern Finland
Dissertations in Education, Humanities, and Theology; 85
ISBN: 978-952-61-2114-7 (nid.)
ISSNL: 1798-5625
ISSN: 1798-5625
ISBN: 978-952-61-2115-4 (PDF)
ISSN: 1798-5633 (PDF)

ABSTRAKTI

Tämä tutkimus käsittelee yhtä englannin kielen varhempien vaiheiden tärkeimmistä muutoksista eli tiukan SVO-sanajärjestyksen kehittymistä. Aiemman tutkimuksen valossa tähän muutokseen ovat vaikuttaneet sekä kielensisäiset että -ulkoiset tekijät. Jälkimmäiseen sisältyvät kontaktit varhaisen skandinaavisen väestön kanssa, mikä on usein mainittu tärkeänä muutoksen syötteenä. Näitä kontakteja pidetään myös yhtenä niistä tekijöistä, jotka ovat vaikuttaneet englannin sijataivutusten rapautumiseen. Tämä puolestaan on muutos, joka on vaikuttanut SVO-järjestyksen asteittaiseen vakiintumiseen. Tutkimuksen ensisijaisena tarkoituksena on arvioida ulkoisten tekijöiden roolia uusien syntaktisten rajoitusten vakiintumisessa keskiajan englannissa. Yleisemmästä näkökulmasta tämä työ arvioi kielikontaktien merkitystä morfosyntaktisten muutosten edistämisessä. Näiden lisäksi tutkimus tarkastelee yhden alan kirjallisuudessa käytetyn kielellisen muutoksen mallin, Johannes Schmidtin ”aaltoteorian”, relevanssia kyseisten muutosten selittämiselle.

Tutkimuksen aineisto koostuu kahdesta jäsennellystä korpuksesta, jotka edustavat muinais- ja keskienglantia: The York-Toronto-Helsinki Corpus of Old English Prose (YCOE) ja The Penn Parsed Corpus of Middle English (PPCME2, toinen painos). Näissä kahdessa korpuksessa on yhteensä lähes kolme miljoonaa sanaa, ja ne sisältävät eri murteilla kirjoitettuja tekstejä. Tutkimuksessa mitataan ja verrataan (S)VO-sanajärjestyksen esiintymistiheyttä valituilla murresektoreilla eri näkökulmista. Näihin kuuluvat lauseet erilaisissa esiintymisympäristöissä, nominaaliset ja pronominaaliset NP-konstituentit sanajärjestyssekvensseissä, ajankohdan ja genren mahdollinen vaikutus tutkituissa tekstiryhmissä sekä sanajärjestysjakaumien eroavuuksien vertailu alkuperäiskielellä kirjoitettujen ja jostakin muusta kielestä käännettyjen tekstien välillä. Aineiston monitahoisen analyysin tavoitteena on muun muassa arvioida jäsenneltyjen diakronisten korpusten käytettävyyttä laajakantoisten kielellisten muutosten jäljittämisessä.

Tutkimustulokset osoittavat, että (S)VO kehittyi nopeimmin murrealueilla, joilla oli eniten kielikontakteja eli Pohjois-Englannissa ja Keski-Englannin itäosissa. Kielikontaktialueilla (S)VO vakiintui erityisesti alisteisissa lauserakenteissa. Tämän lisäksi pronominien ja substantiivien normalisoitu esiintymistaajuus oli korkein pohjoisimmissa murteissa. Nämä tutkimustulokset viittaavat siihen, että kielenulkoiset

pikemminkin kuin -sisäiset tekijät motivoivat rakenteellista muutosta. Myös geneettisten ja arkeologisten tutkimusten tulokset todistavat selvästi skandinaavisesta alueesta, joka ajallisesti ulottui niin kutsutun Viikinkiajan ulkopuolelle. Morfologisen yksinkertaistumisen ja uuden sanajärjestyksen kehittymisen yhteys on hyvin todennäköinen, sillä (S)VO:n esiintyminen yleisimmin ja tasaisimmin toistui tuloksissa erityisesti pohjoisessa. Pohjoisen hallitseva asema murrekirjossa viittaa myös keskeisen alueen olemassaoloon, jolta kielelliset innovaatiot saivat alkunsa aaltoteorian esittämällä tavalla. Keskiajan Englannin sosiopoliittista taustaa vasten tarkasteltuna tutkimuksen tulokset vahvistavat pohjois-etelä -jakauman olemassaoloa. Pohjoinen alue muodosti suotuisat puitteet erilaisille muutoksille, jotka koskivat myös kielellisiä muutoksia.

Avainsanat: englannin kieli, sanajärjestys, taivutusmorfologia, kielikontaktit, murteet, variaatio, korpustutkimus

Acknowledgments

My investigation into contact-induced changes in Early English began with my MA thesis, in which the impact of Old Norse on Old and Middle English vocabulary was addressed and its extent assessed. The present study is a follow-up to that early research, with the focus placed on morphosyntactic changes resulting from the contact at issue. The ideas for both studies were born in Joensuu. Throughout the years of my work, there have been numerous people and institutions that ensured my academic stride and to whom I owe a great debt of gratitude.

Firstly, I would like to thank my supervisors, prof. Markku Filppula and prof. Stefan Werner. I am immensely grateful to prof. Filppula for instilling inspiration and maintaining my passion for pursuing the topic as well as for his unfailing support and patience. The hours we spent discussing problems of linguistic study are to be cherished. He is a great mentor and an understanding boss. Successful completion of my study would have been harder to attain without the continued assistance of my second supervisor, prof. Stefan Werner, who unrelentingly guided me through the intricacies of computer language and statistical calculations.

Next, I would like to thank the external pre-examiners of this thesis, Professor Emeritus D. Gary Miller, University of Florida, and Docent Leena Kahlas-Tarkka, University of Helsinki, for agreeing to read my work and providing me with numerous critical and the same time valuable comments. With their expertise, they have in many ways contributed to the final make-up of the manuscript and provided me with pertinent suggestions on necessary modifications of the content. I am also both grateful and honoured that Docent Kahlas-Tarkka has agreed to act as the opponent at my defence.

My deep gratitude goes to the professors and PhD students in Langnet, The Finnish Graduate School in Language Studies, especially those engaged in the Language Variation and Change programme, as well as the people in the research teams of the Global English (GlobE) and Changing English (ChangE) Consortia for their incisive feedback and valuable suggestions during seminars and beyond. I would also like to acknowledge the assistance of the staff and administration of the University of Eastern Finland. They ensured that I, a non-Finn, would not meander too long through the various procedures and sometimes perplexingly complicated paths of document handling.

The financial support I have received during my work has been crucially important. I wish to thank the Philosophical Faculty of the University of Eastern Finland, Langnet, the GlobE and ChangE Consortia, as well as the Ella and Georg Ehrnrooth Foundation, for making this research possible.

I am everlastingly grateful to my work colleagues and friends at the university and off the academic setting, especially to my wonderful Finnish crew members: Lea Meriläinen, Sanna Hillberg, Paula Rautionaho, and Heli Paulasto, who have made me feel a part of the university family and have always offered their helping hand when

confronted with the complexities of the Finnish language. A big thank you goes to Marta Choroszewicz and Attila Krizsán. Their friendship and moral support have been invaluable throughout the years. They were also the ones who provided great solutions when things came to a deadlock and who would find me when getting lost in unfamiliar surroundings, literally!

I am immensely indebted to my friends and closest ones back in Poland, who have been by my side through thick and thin. Special thanks go to my lifelong friends Ula Kijewska and Radosław Potorski, who have made sure I would remember that life away from books also matters. The most heartfelt gratitude goes to my dearest parents, for their unconditional love and undying faith in my endeavours. You helped me get where I am today and you never doubted my success. Finally, I thank my partner, Pasi, for being an endless source of motivation and confidence and, most importantly, for keeping me vertical.

Joensuu, April 2016

Izabela Czerniak

Contents

FIGURES

TABLES

1 Introduction

The present study examines the emergence of strict SVO order in English. It is the structural framework which constitutes the major determinant of grammatical relations in the English clause and which has successfully thrived to this day. The research focuses on the earliest stages of development of the language, the closing period of Old English until the end of the Middle English period. During that time, English lost much of its inflectional load and began to turn away from its relative syntactic freedom in favour of a more rigid order of constituents. The clause-initial noun phrase, left of the verb, would come to function more frequently as a subject, whereas the one placed postverbally would act as an object, whether nominal or pronominal. The verb phrase, too, began to exhibit more continuous than discontinuous tendencies. Finally, a typically Germanic verb-second constraint became increasingly restricted, which distanced English from the language family matrix.

Although morphological simplification was one of several contributing factors in the change to the prevalent SVO, the loss of cases on NPs has been regarded as particularly instrumental in the postverbal placement of objects. What is more, not only has the close relationship between the case and SOV/SVO order been typologically attested in other languages but the earliest instances of SVO in English, too, have been confirmed to involve contexts where subjects needed to be differentiated from objects. The current research, thus, aims to estimate to what extent the loss of (case) inflections influenced the subsequent syntactic reanalysis. Since the timing of the onset of morphosyntactic change in English coincides with the emergence of linguistic effects of the early Anglo-Scandinavian language contacts, it is pertinent to assess the role that these contacts played in the structural shift discussed here. In the literature, the loss of case inflections, among other linguistic changes, has often been attributed to this particular contact situation (see the discussion in chapter 4). The contact scenario apt for this kind of externally induced change would be that of convergence, whereby the emphasis is placed on increasing structural similarities or, conversely, eliminating contrasts between the two languages.

The arrival of the early Scandinavian invaders has been universally associated with the onset of the so-called Viking Age (*Vikingetiden*) starting at the end of the eighth century (Geipel 1971: 30). After the initial bloodshed and pillage, assimilation followed and the two populations eventually settled down side by side. Long-standing bilingualism flourished, as the successive Scandinavian generations integrated into the Anglo-Saxon society. The amalgamation of the two populations, additionally stimulated by loosening ties with the North Germanic homelands of the Vikings, meant that there was less and less motivation to keep the Norse language alive. Indeed, sources mention the steady decay of Norse during the first half of ME (e.g. Townend 2006: 66, 68-9; cf. Geipel 1971: 56-8). However, the Scandinavian contribution did not weaken along with the withering Norse vernacular. From the beginning of the Middle English period there were Scandinavian words transpiring in the texts. Interestingly,

the instances were not confined to content words only. Closed-class items such as pronouns, conjunctions and prepositions were borrowed as well (e.g. Jespersen 1935: 71). It is also likely that apart from the morphological simplification mentioned above, some of the structural tendencies were, at that time, adopted from the Norse language, e.g. the non-clitic status of pronouns, omission of *that* as well as the increased propensity for surface VO. This study explores the extent of Scandinavian influence which could have gone as far as inducing a fully-fledged syntactic shift. Although strict SVO word order in English eventually developed as a result of a combination of various secondary processes, it seems feasible to see the contact-induced loss of inflections as an important impulse towards the syntactic change. Increased opacity of syntactic relations and the apparent syntactic discomfort experienced by speakers and writers could be the relevant factors, especially at the earlier stages of the development of the language. At that point, a rapid receding[1] of a complex inflectional paradigm needed to be supplanted with other mechanism(s) of expressing grammatical relations. The most straightforward and unambiguous choice would be to promote a familiar, already available option, viz. SVO word order.

Research on word order in English has been conducted for well over a century. The first significant attempt at explaining the intricacies of early English structures was made by C. A. Smith in 1893. He was looking for what he called "the syntactic norm" (Smith 1893: 212), whereby the word order would be determined by syntactic factors rather than by rhetorical or euphonic aspects. Since then, various proposals have been put forth. There have been synchronic accounts with a non-generative bend, to which Smith's study belonged. Later accounts of the synchronic-non-generative type include, among others, the work of B. Mitchell (1964, 1985), where the role of the subject in the structural framework is emphasised. Next, there have been synchronic explanations with generative underpinnings. Studies by W. Koopman (1985, 1990, 1992), S. Pintzuk and A. Kroch (1985, 1989) make good examples, all relying on verb-finality as the underlying OE structure and allowing for the V2 rule.

On the other side of the spectrum, there are the diachronic, non-generative explanations. Notable instances are provided by the works of B. M. H. Strang (1970), T. Vennemann (1974, 1984), M. C. Bean (1983) and V. Kohonen (1978), who have explored numerous factors conditioning word order preferences, viz. the "weight" of elements (Strang), the importance of the topic in the marked sentences (Vennemann) and the difference between matrix and subordinate clauses (Kohonen). The study by Bean (1983) tests multiple ordering principles simultaneously – T+V, V1, V2 and V3[2]. Lastly, there are diachronic scenarios based on the generative point of view. Research by van Kemenade (1987) can be added to that group. Her work on the role of OE and ME pronouns in determining word order patterns has been both significant and very informative. Other studies which need to be mentioned are that of F. Colman (1988), where the constituent weight is revisited and refined, along with the proposals of Stockwell and Minkova (1990, 1991, 1992), in which subordinate clauses play a vital

[1] As accelerated by external pressures.
[2] T+V: topic before the verb, V1 (V2, V3): verb first (second, third) position in a clause.

part in the establishing of the new V3 order[3]. Recent attempts at describing the syntactic change in early English have involved most of the aspects explored earlier but augmented with the input from the ever-developing branch of corpus linguistics and accompanied with solutions provided by both language internal and external forces. The current research is one of these attempts.

Admittedly, investigating language contact as a sole impulse leading to a major syntactic change of the type witnessed in the history of English would have to be considered too simplistic because pressures found inside and outside languages tend to work in tandem. Nonetheless, this study aims to show that the role of external pressures in the structural shift in English has been underestimated and considering these pressures as a mere accelerator of the change would be inaccurate as well. In this study, the extent of the early Scandinavian influence will be gauged by a longitudinal comparative study of 'new' word order distributions based on two parsed corpora representing Old and Middle English: YCOE: *The York-Toronto-Helsinki Corpus of Old English Prose*; and PPCME2: *The Penn Parsed Corpus of Middle Englis*h (Second Edition). The rate of the spread of the 'new' (S)VO word order will be investigated in particular dialect sectors from various comparative perspectives relevant to the study. These perspectives include the difference in clausal environment, the rate of nominal against pronominal constituents of NPs in word order sequences, the importance of date, as well as the genre of texts included in the sectors. Yet another aspect explored here is word order distributions in texts representing foreign translations as against the native originals. The significance of the present study lies in its multidimensional approach, whereby the many factors influencing the structural layout of the clause and its development in early English are investigated simultaneously.

The theoretical framework chosen to validate the hypothesis that the emergence of SVO can be traced back to a common point of innovation, which subsequently led to the structural reshuffle, is Johannes Schmidt's wave theory (*Wellentheorie*). This theory emphasises the importance of a focal area, the location where externally motivated impulses converged to promote new linguistic tendencies in such a way that they could further spread at more or less steady tempos to other, peripheral regions. Statistical tools such as the coefficient of variation are employed in this study to measure the level of condensation of the investigated feature in data sets. The input offered by archaeology and genetic studies is also used in the analysis of the social history of the areas affected.

The discussion in this study proceeds as follows. Chapter 2 presents the sociohistorical background to the morphosyntactic change at issue here. It starts with a description of the populations that occupied the British Isles before the onset of the Viking Era. The following section outlines the abrupt and violent landing of the North-Germanic tribes in the British Isles with the subsequent assimilation and acculturation in the early English society. This section also includes a discussion of the Scandinavian incursions prior to these large-scale invasions. Having provided an outline of the Scandinavian arrival and settlement onto the English soil, an account

[3] Outline of word order studies is based on Denison's summaries in *English Historical Syntax* (1993).

of the dividing of the country into two, almost separately governed territories, viz. the North and the South, will be given next. The chapter ends with a description of the socio-linguistic situation in the post-invasionist England. Chapter 3 provides the theoretical essentials on causes, mechanisms and factors behind the changes of word order. Here, the role of morphological simplification in these changes is identified. The subsequent section deals with motivations for the changes in morphosyntax, which encompass both internal and external pressures. The chapter ends with the outline of models of progression of change through the socio-geographic space. Chapter 4, then, gives an account of the morphosyntactic changes in early English alone, with some reflections on the potential causes of them. Before presenting the chronology of the morphological simplification and expanding on the likely scenarios on the emergence of the SVO order, the fundamentals concerning the stage of the English language preceding the change will be given. The following sections tackle particular instances of foreign influence on early English. Special attention is paid to the Scandinavian input. The chapter closes with a proposal concerning the contact scenario relevant to the shift presently discussed. It also presents the wave theory as the most suitable model of the spread of linguistic change for the purpose of this study. Chapter 5 focuses on the research questions and aims, and presents the databases used in this study as well as the methodology used in the analysis of the data. An important part of the chapter is devoted to the delimitation of the aspects considered with respect to the emerging word order, most of which are conditioned by the nature of the corpora. Chapter 6 consists of the analysis of the data. The chapter begins with a diachronic view, tracking the route of the emerging grammatically functional word order from Old through Middle English. It then moves up to a synchronic study of Middle English alone. The following sections focus on finding out the extent of the Scandinavian impact in the light of the wave model. The results are consolidated and interpreted in chapter 7. The final discussion revolves around the contribution of this study to the research on early English syntax and includes possible directions for further investigation.

2 Socio-historical Background

This chapter presents the socio-historical background to the language changes investigated in the present study. Although the primary focus has been placed on the medieval contacts between the English and Scandinavian populations, the outline covers a period much wider than the time frame usually adopted for providing an accurate storyline, beginning before the Dark Ages and ending with the close of the fifteenth century. It is vital to move slightly further back in time in order to understand the reasons for as well as the outcomes of the events which brought about the most significant shift at the early stages of development of the English language.

A few points concerning the appropriate nomenclature ought to be clarified before proceeding with the socio-historical account. The term 'Scandinavian' will be used throughout to denote any population group of the North Germanic tribes, be it those referring to the people of Norway, Denmark, Sweden or Iceland. The term includes, therefore, the groups arriving before the Dark Ages as well as the raging warriors of the late eighth through tenth centuries, also encompassing the settlement of the less bellicose Scandinavian groups who followed suit (cf. Fisher 1973: 208, 211). In addition, the name 'Viking', which could either refer to[4] a warrior (*wiggend*), a pirate (*vikingr*), a person (or people) who lived by the *vik* (ON: bay), relay oarsmen (*vika*) or an armed camp (OE: *wic*)[5], will point to the early Scandinavian warriors exclusively, with a clear distinction between Danish and Norwegian fractions where necessary (cf. Rowland 1993: 66). Finally, the term 'Norse' will be used with an adjectival meaning of 'old Scandinavian', designating the early North Germanic language in order to avoid confusion between 'Norse' meaning either the early Norwegian population only (cf. Collingwood 1993: 59, Higham 1993: 173) or the entire early Scandinavian race.

2.1 PRE-VIKING BRITAIN

Before the areas around the North Sea became a playground for part of the mass migration of Germanic tribes, including the arrival in Britain of the Anglo-Saxons and later the early Scandinavians, the British Isles had already housed three distinct population groups, the Picts, the Celts (i.e. Britons), and the Romans. Each of these groups left an indelible demographic as well as linguistic imprint on the Isles. The current section aims to provide a short description of the presence of these distinct populations. Their impact on the subsequent socio-linguistic make-up of the Isle has been well attested. The section ends with a rather brief account of the incursions

[4] Or derive from.

[5] Sources: Collingwood (1993: 60-63), Logan (1983: 28), Roesdahl (1991: 9), Richards (2005: 4). Richards also remarks that both OE and ON forms were suggested as parallel developments from a Germanic word 'to withdraw, leave or depart (2005: 4).

and successive settlement of the Anglo-Saxons. The arrival of these Germanic tribes makes the bedrock of what has become the English nation and language.

Although the Picts, the original population to dwell in the Isles[6], did not play a particularly influential role in the future development of the English language, their presence needs to be accounted for in order to accurately assess the layout of the first inhabitants of the Isles. These "supposedly painted, aboriginal tribes of northern Scotland", as described by Oppenheimer (2006: 71), possibly constituted a conglomerate of the highland tribes[7], arriving as a consequence of the ever expanding Roman Empire (e.g. Foster 2004 quoted in Miles 2006: 181). The extent of their lands has been identified by the symbol stones (e.g. Sykes 2007: 218) as well as by the characteristic 'Pit-' place names, e.g. Pitolchry or Pitmadden (Miles 2006: 181). Bede clearly recognised the Picts as a separate nation, apart from the Britons and the English. He claimed that their homeland would be found in Scythia, which, at some point, had been connected to the ancestral Scandinavian homeland mentioned in the Danish and Icelandic sagas (e.g. Cunliffe 2004: 317, cf. Herodotus as quoted in Miles 2006: 105). The early Scandinavian-Pictish ties are, indeed, intriguing. While the native soil of the Picts might still be an unidentified territory, the incorporation of an early North Germanic tint into the Pictish lore has been confirmed, as exemplified by the indigenous (Pictish) monuments in Scotland with clear Scandinavian overtones (e.g. Cramp 1982: 12-13, for genetic evidence see Sykes 2007: 323). The sources also speak of the early Scandinavian take-over of the Picts along with the lands they occupied. The northern incomers were, among others, responsible for killing Eoganan (circa 839), the last Pictish king, as recorded in the Irish annals (Higham 1993: 228, cf. Williams *et al.* 1991: xxxii). The 'Painted Ones' became a lost people, although their genes have survived in the blood of the Scottish (Miles 2006: 181). Indeed, the late ninth century stood witness to the amalgamation of the kingdoms of the Scots and the Picts (Blair 1977: 44-5). Linguistically, they seem to have been supplanted by Scottish Gaelic (at least in western Scotland) and eventually by English (Oppenheimer 2006: 71).

The mention of Gaelic brings forth the second population group to occupy the Isles before the arrival of the Anglo-Saxons, the Celts (i.e. Britons). The earliest mentions of Britain with reference to the Celtic population comes from the *Massalliot Periplus*[8], written circa 600 BC (Miles 2006: 107). The conventional academic view, on the other hand, holds that the Celtic cultural history in the Isles began no earlier than 300 BC (Oppenheimer 2006: 87). The correct answer seems to be leaning more on the former time estimate. The archaeological evidence from Later Bronze Age Britain shows that the Isle was already quite populated and the land considerably cultivated (e.g. Miles 2006: 109, 135). The insular mode of life had kept the Celts slightly at odds with what was regarded as civilised. Otherwise, they might have seemed relatively backward as compared with the progress enjoyed by their Continental relatives at the time the Romans landed in the Isle. When Caesar came to Britain in 55 and 54 BC, he was sur-

---

[6] As verified largely by evidence from aerial photography, fieldwork and excavation (Miles 2006: 182).
[7] Taking up the area of Scotland specifically north of the Firths of Forth and Clyde (Fisher 1973: 40).
[8] The description of a voyage from Marseilles. The account of *Periplus* survives in the *Ora Maritima* of Avienus of the 4th c. AD (Miles 2006: 107f).

prised to find the native warriors using chariots in the battlefield, a relic of the past (Miles 2006: 110). Caesar's coming to Britain paved the way for the full-scale invasion, which was to come a hundred years later. He managed to submit the British tribal chiefs and made them pay tribute. Expectedly too, he put puppet kings onto the local thrones so that by the time the Romans properly invaded Britain, most of the British population was already under their control. If there was any resistance, it would soon be overcome. As for the Celtic way of life under the Roman rule, the natives were not so powerfully forced into the strict following of the occupant laws and customs. So long as they observed the basic facts of who was the master, the Britons enjoyed their 'freedom'. The south-eastern fraction of British aristocracy, in fact, openly adhered to the more civilised ways of their Roman supervisors. The trade relations with the closest parts of Gaul definitely enabled it (Sykes 2007: 293).

Linguistically, the Celts provided two important contributions to the language of the early English, their native (British) Celtic and their rendition of Latin. The latter was introduced by the Romans largely through schooling, a common practice in the provinces (Clackson and Horrocks 2007: 230-231, Hoad 2006: 10, 30). Latin would be the only written medium in Britain at the arrival of the Germanic tribes (e.g. Schrijver 2002: 89).

The essentials on the Romans, who were the last ones to attempt to people and settle in the British Isles before the so-called *adventus Saxonum*[9] in the mid fifth century AD, have already been presented in the previous paragraph. The Romans made Britain part of their Empire for about four centuries (Blair 1977: 1). During that period only half of the Isle was in their charge, the rest was a barbarian territory (Miles 2006: 130). The occupation of the land and the relationship with the natives was relatively harmonious though not completely diplomatic. Yet, as suggested by the evidence excavated, the growing number of native settlements, of the fields as well as the amount of the forest clearing, reveals the expanding population under the Roman rule. However, only the native aristocracy could enjoy unrestrained wealth of the Civilisation (Miles 2006: 131, 151). The Romans gave Britain decent sites – the foundations to the future towns, as well as fortifications, with a fine network of roads, some of which serve their purpose even to this day (e.g. Carus-Wilson 1958: 211). They established new ways of producing goods and consuming them. Overall, Britain during the Roman times presented a multicultural blend of people, customs and religions (Miles 2006: 148). Traditionally, the end of Roman Britain has been dated to the year 410 and onwards (Miles 2006: 161).

As to the beginnings of the English race in the Isle, there is the continuously passed down account of the Britons inviting the Anglo-Saxons to help them fight off the rowdy Picts/Scots in the north (e.g. Fisher 1973: 15-25, Rowland 1994: 11). The very same Germanic tribes, who promptly saw the British soil ripe and fit to settle, soon became the enemies, as they turned against the natives themselves. The sources mention initially the coming of small Germanic communities, crossing the English Channel, which soon turned into a large-scale colonisation of Britain by the Angles, Saxons, Jutes and also Frisians from the beginning of the fifth century (e.g. Roberge

[9] From Filppula (2010: 432-3, also Miles 2007: 164).

2010: 419). The conquest was a slow and drawn-out process, divided into two phases. At first, hard fighting and battles predominated. The result was only killing, escaping or enslaving of the native inhabitants. The Germanic groups were originally assembled into independent war bands which, subsequently, formed larger groupings. The second phase of the invasion was conducted by military leaders who had more politically oriented goals in their minds.

The beginning of the seventh century saw the formation of the *Heptarchy* (Pollington 1989: 95), the seven Anglo-Saxon kingdoms[10]. Out of these kingdoms, Northumbria, Mercia and Wessex soon became dominant and continued to muscle for supremacy up until the coming of the Vikings. Fisher (1973: 89) notes that the balance of power between the three Anglo-Saxon realms changed so often that none of their kings could be regarded as the proper *bretwalda*[11]. Initially, it appeared that Northumbria would be the one to exercise authority over the rest of the Heptarchy. The fact that the far north, the land of the Picts, was not causing any disturbance during the seventh century allowed Northumbria to focus on the battles for control with the southern English, especially with the Mercians (Fisher 1973: 87 and 141-2). However, the kingdom of the northern English would lose its definitive say in the matters of the Anglo-Saxon realms. The main reason for the eclipse of Northumbrian power (circa eighth century) can be found in its internal disputes, in the continuous disagreements between the (rival) members of the royal family (Fisher 1973: 110, 121 and 148). The Mercian kingdom would take over the sceptre of authority, as the glory days of Wessex were yet to come (Fisher 1973: 156). Nonetheless, these were the changing winds of politics only. In terms of scholarship and arts, Northumbria would remain the unquestioned authority for quite a long while (Fisher 1973: 161). The greatest specimens of learning and craftsmanship were created in Northumbrian monasteries (Fisher 1973: 156). It is the northern Wearmouth-Jarrow campus, for instance, that produced the top scholar, Bede. There was also an exquisitely furnished library found in these walls, established by Benedict Biscop, who brought books from the Mediterranean lands. The stock was later enriched by Abbot Ceolfrith (Fisher 1973: 156). York was another location in the north where scholarship flourished, with a library of size and stock easily comparable to that of Bede (Fisher 1973: 157). People came from as far as Frisia to York for study (Fisher 1973: 188). What is more, the northern capital exemplified another aspect in which Northumbria outshined other Anglo-Saxon kingdoms – architecture.

Mercian supremacy was distinguished by the reign of two kings. The first *Aethelbald rex Britannia*, the king of Britain according to the charter of 736, was murdered by his bodyguard in 757. The aftermath of his death was the impending the civil

---

[10] The Angles split up and founded three separate realms. The first was Northumbria, which was formed out of former Bernicia and Deira, with the capital established in York. The second was East Anglia and the third, Mercia, with Lichfield and Tamworth as its main towns. As for Saxons, the eastern fraction settled around Colchester, creating Essex. The South Saxons set up the kingdom of Sussex and the West Saxons, expectedly, Wessex, with the capital in Winchester. Finally, the Jutes took the south-eastern part of Kent and the Isle of Wight. Canterbury was their centre (e.g. Miles 2006: 164, Sykes 2007: 306, Fisher 1973: 108-9, Wales 2006: 34).

[11] OE: 'ruler of Britain' – the overlord of the (southern) English in the period before the unification of England (Williams *et al.* 1991: viii).

war (Fisher 1973: 163). Offa was the second, named *rex (totius) Anglorum (patriae)*, which was quite an achievement at that time. Unlike Aethelbald, the ruler over the Southern English, he was (aspiring to be) the king over all the kingdoms, including Northumbria (Fisher 1973: 168, cf. Loyn 1994: 23). Offa's reign, among others, established the great (court) tradition of Mercian charter scribes. So much so that the Wessex kings would turn to Mercian (translation) expertise for help (Fisher 1972: 229, Blair 1977: 351). The political dominance and contacts with the Continent made the Mercian culture unparalleled (Fisher 1972: 170). Yet, just as the murder of Aethelbald induced instability in the land of Mercians, the death of Offa in 796 marked the end of their power within the Heptarchy (Fisher 1972: 199). The kingdom's fate was decided by the arrival and settlement of early Scandinavians. Along with it fell the fine Mercian literary tradition (Knowles 1997: 37).

By the ninth century, Wessex slowly began to come out of dependence induced during the reign of Offa and grew to be the strongest of the English kingdoms (Fisher 1973: 198, Blair 1977: 53, Knowles 1997: 27). Of all its leaders Alfred was the only one to match (or surpass) the greatness of Offa (Blair 1977: 53). It was in Wessex, too, that the Anglo-Saxon learning was revived (Knowles 1997: 38), notably due to Alfred's initiative, and that the English (rather than Latin) writings blossomed (Blair 1977: 329 and 350-1). Alfred was quite aware of how much English administration would benefit if the leading people of each region could read and understand king's orders themselves through letters and codices instead of waiting for the clerks to translate them (Pollington 1989: 77). Suffice to say that the West Saxon dialect, eventually, became the dominant variety of English (10th c.). Thus, quite expectedly, the majority of surviving texts were written in the West Saxon vernacular. The linguistic differences between the early English dialects would be sustained, nonetheless, even with potential levelling occurring initially (Trudgill 1986: 288). The differentiating dialectal aspects, it seems, would have survived the invasions (Wales 2006: 34) if not strengthened by unbroken liaisons with Scandinavia (Knowles 1997: 34). Further, during Alfred's reign the (legislatively validated) foundations of England as a sovereign nation were laid (Fisher 1973: 231). Finally, he was the one who had to deal with the new external threat – the coming of the Vikings. What remained of England after the emergence of the Scandinavian Danelaw was controlled predominantly by the (old) kingdom of Wessex (Knowles 1997: 35) and protected by the well constructed system of fortifications[12], devised, again, on Alfred's orders (Blair 1977: 76). Yet, a century or so had to pass before the tight Scandinavian grip loosened and England could be united again.

Lastly, along with the establishment of new communities, the new religion was introduced and the Anglo-Saxons would yield to Christianity, though not all derived from the same source. The new Christian faith was, in reality, introduced into Britain already in the second century but one can track its progress only from the fourth century onwards (Fisher 1973: 55). However, it took a while before the religion became properly ensconced in Britain and, as archaeological evidence along with onomastics evidence shows, paganism was commonplace well into the Anglo-Saxon times (Fisher

[12] With the *Burghal Hidage* serving as evidence (Blair 1977: 76, Pollington 1989: 153).

1973: 63). Christianity would be more welcomed in (southern) Ireland, through the impact of British and Gaulish traders (Fisher 1973: 58) and it is due to the Irish monasticism (St. Columba's Iona) that Christianity began to flourish in the North (Fisher 1973: 62). The literacy and the distinctive Anglo-Saxon handwriting of the eighth century was also brought in with Irish monks (Blair 1977: 313, Knowles 1997: 25). The South, on the other hand, was influenced by the mission from Rome issued by Pope Gregory. The mission was initiated in Kent in 597 under the leadership of Augustine (Whitelock 1974: 156)[13]. The Anglo-Saxon Christianity thrived and so were the monasteries along with the so-called pseudo-monasteries[14], governed by laypeople (Fisher 1973: 172). The unstable political situation during the Viking raids towards the end of the eighth century and the subsequent dislocation and destruction could have created some impediments to the normal work of the Church in England. Yet, there is no evidence available to support that claim (Fisher 1973: 203). However, it is evident that the arrival and settlement of the early Scandinavians created a reset in socio-political relations (Fisher 1973: 43, 113, 124, 141, 196 and 295). The Viking assaults persisted up to the point when the Scandinavian king Cnut (anglicised as Canute) took the English throne (Williams *et al.* 1991: xiv). The following section covers the essentials on the period in England commonly referred to as the Viking Age.

2.2 SCANDINAVIAN INVASIONS

Before proving the account of the Viking Era, the issue of the earlier Scandinavian presence in the British Isles needs to be addressed. It has been frequently brought to light in recent years that the North Germanic tribes were not only very much familiar with the Isles prior to their attested entrances but also that they were already present there, with, at least, small-scale colonies and trading ventures[15].

The first recorded violent landing of the Scandinavian warriors in 789, when the Vikings "did not take kindly to being admonished" by the local king's reeve, Beaduheard (Blair 1976: 222, also Geipel 1971: 32), was most probably not the first step that the Scandinavians made on the British shores. At least we can speak of fairly

[13] Halfway through the 8th c. there were as many as four dioceses in the North, the above mentioned York and Hexham, also Lindisfarne, which was soon looted by the Vikings and finally Whithorn. The South had twelve sees, expectedly located in the kingdoms' capitals such as Selsey, Winchester, Sherborne, London, Dommoc and Lichfield. The Kentish Canterbury would be the seat of the archbishop. Other southern sees were founded in Rochester, North Elmham, Worcester, Hereford and Lindsey (Blair 1977: 143). They were the centres of religious guidance and intellectual attainment.

[14] These were often frowned upon and, as noted by Godfrey (1962), they were no more than "places of refuge for renegade monks and for members of society who sought to evade all responsibility" (Godfrey 1962: 165).

[15] Earlier Saxon presence in the Isles has been highlighted as well. They were already in residence during the Roman times, recurrently employed as mercenaries or legionaries (the so-called *foederati*). By the 3rd century, the regular Roman legions would be fairly Germanised (Fisher 1973: 15, Oppenheimer 2006: 308). There are also references to the famous *Saxon Shore* (*litus Saxonicum*) (e.g. Fisher 1973: 14-16, Blair 1984: 16 although cf. 1977 4f, Collingwood 1993: 45, Miles 2007: 159). It was a set of Germanic military sites (forts) or civilian settlements put up along the easternmost coasts of Roman Britain (Oppenheimer 2006: 311). The debate continues over whether the location was settled and defended by the Saxons or set up against them (cf. Higham 1993: 29, 49).

well developed Anglo-Scandinavian diplomatic relations taking place long before the Vikings began to raid and plunder the Isle(s). On one hand, scholars speak of the alliance between the Anglian kings of England and their cousins[16] in Denmark valid already since the sixth century. On the other, there is the archaeologically attested historical contact between Norway and Britain, through links with the coastal districts of England during the fifth and sixth centuries as well as with the north-western parts of the Isle beginning from the seventh century AD (Oppenheimer 2006: 386, Fisher 1973: 175, Roesdahl 1991: 189). All these territories were, in due course, included in the Scandinavian impact and subsequent settlement zone[17]. Archaeologists speak about "an intermittent two-way flow of cultural influences – and, so presumably, of human beings – between Britain and the lands to her north and east that began long before the final rupture of the islands from the parent continent" (Geipel 1971: 28, also Crawford 2003: 41, Collingwood 1993: 46, Logan 1983: 41). The recent genetic studies support that claim. For example, Oppenheimer (2006) shows that the Vikings did not import all Scandinavian gene markers to the British Isles and that there were repeated earlier inflows (2006: 385-6). The Scandinavian raiders of the eighth to tenth centuries were, according to Oppenheimer, landing in their "former haunts" (2006: 414-15)[18]. In addition, his study confirms the presence of Danish gene predominantly in the areas under the Danelaw, especially York and the Wash (Oppenheimer 2006: 390-4, cf. Sykes 2007: 333, 337). It also shows that the Norwegian markers are more widespread, not so localised as their Danish counterparts, and that the Norwegian influence in northern Britain is clearly older than generally assumed (Oppenheimer 2006: 400).

To recognise these early incursions means changing the manner in which we should interpret the main invasions of the early North Germanic tribes. The aftermath of these events should be re-examined as well, viz. the pace of integration of 'newcomers' into the native society and the formation of regional (including linguistic) differences arising on that account (e.g. Oppenheimer 2006: 264-5). Acknowledging the presence of the Scandinavians prior to the large-scale landings creates the picture of the natives being subsequently shocked not so much at the arrivals of the frighteningly foreign tribes but rather at the great abundance of people who were violent and perhaps not exactly of the same faith (Fisher 1973: 62, Morris 1982: 78-9, cf. Logan 1983: 36).

Alcuin's lament of 793, in his letter written to Ethelred, king of Northumbria, shortly after the raid on Lindisfarne, sheds some light on how the countrymen took to the early Scandinavians. He was appalled at the fact that the natives were aligning with the Vikings by turning away from Christianity and by imitating their style of clothing as well as hairdos. In Alcuin's frame of mind, the natives' reaction to the recent Scandinavian horror ought to have been completely different: "(...) Are

---

[16] Scandinavian (Swedish) roots of (East-)Anglian population are mentioned by Fisher (1973: 108 and 116).

[17] Essex proves to be a pertinent case in point. Although this Saxon kingdom was incorporated into the Danelaw, it was never colonised by the Danes (Oppenheimer 2006: 390).

[18] Both male and female Neolithic genetic link with Norway has been identified in the extreme north of Scotland and its neighbouring islands (Orkney, Shetland, and the Western Isles). Neolithic genetic connections pointing to southern Scandinavia and Denmark are located in Eastern England (Oppenheimer 2006: 415).

you not terrified of those whose hairstyle you wanted to have?" (quoted in Frank 1989: 53 also Rowland 1994: 66-7). The newcomers were everything but unfamiliar to the oppressed. The volatile manner of these incursions, as noted by Sawyer (1971) with respect to the beginning of the Viking Age, was nothing more but a continuation of earlier habits (Sawyer quoted in Roesdahl 1991: 189-190, also Fisher 1973: 211, Collingwood 1993: 63-4).

Ultimately there was a difference between cruel and wild Scandinavian pirates, including their later, more civilised rendition, the famous Jóms-vikings[19] (Fisher 1973: 309, Collingwood 1993: 85 and 162, Richards 2005: 4) and the rest of the Scandinavian folk, some of whom were fishermen, farmers and/or merchants (especially Swedes) (Fisher 1973: 208-9 and 212). The last two were typical Scandinavians of the pre-Viking period (Fisher 1973: 211, Geipel 1971: 30). The outcome of the violent landing of 789, when Beaduheard was axed to death by the Viking band, is very much symptomatic of how expected the northern peaceable traders were on the Wessex coast and how unlucky the reeve was to meet the other type of seafarer (cf. Geipel 1971: 32). King Alfred, in due course, could clearly distinguish between the two, as he welcomed the royalties and merchants from Norway while concurrently fighting off the Viking gangs (Roesdahl 1991: 191 and 196). It seems hardly a coincidence that the Angles (and Jutes) would form alliances with Scandinavia, with their homes becoming subsequently a specific target for the North Germanic settlements after the main invasion. Archaeological and gene studies[20] both point to clear links between the two population groups which go beyond the Viking period (Hines 1984: 286-301, Oppenheimer 2006: 413). It becomes evident, therefore, why particular tribes chose particular locations (e.g. Collingwood 1993: 66, 71) and why we are dealing with repeated inflows of populations, disembarking at recognisable shores and not exploring unknown areas (e.g. Oppenheimer 2006: 414-15). What is more, it makes more sense why, for instance, the process of acculturation (assimilation) of Scandinavians took a relatively short time (e.g. Logan 1991: 172). Finally, the earlier incursions as well as the (Anglian) links with Scandinavia might have had their share in the fact that dialects of Old English were already diverging so much from each other in England soon after the 'official' Anglo-Saxon landing (Oppenheimer 2006: 304, also Wales 2006: 33).

As regards the Scandinavian first recorded incursions on the English shores, Dorset (789 AD)[21], Lindisfarne (793) and Jarrow (794), hardly anyone was aware what these violent incidents were to signify (Logan 1983: 40). Interestingly, the Scandinavian motivation for incursion and conquest of the British Isles followed that of the Anglo-Saxons and equally matched the reasons for the great Germanic *Völkerwanderung*[22].

---

[19] I.e. community of professional Viking soldiers who had "stringent set of customs of their own" (Collingwood 1993: 162). They resided in Jómsborg, the most probable site of which is located at the mouth of river Oder (Fisher 1973: 309).

[20] Oppenheimer (2006) claims that the Angles and Jutes were more Scandinavian (than Saxon) culturally and linguistically. There are clear genetic matches to the Danish Peninsula and Sweden (Oppenheimer 2006: 413).

[21] The AS Chronicle gives the year 787, describing the arrival of Northmen in (Portland) Dorset. Logan (1983) corrects it to 789. According to him, the Chronicle "at that point is two years out of synchronisation" (1983: 38).

[22] See Roberge (2010) or Miles (2006: 161) for details.

Roberge (2010) lists as many as five causes for Germanic intrusive migratory movements. Among them, there are the all familiar overpopulation and the competition for resources in the homeland, which would have inevitably occurred because of the former. He also mentions the role of pressures and perspectives coming from outside. One of those was untold wealth obtainable from the foreign (southern) lands; the other, the consequences and opportunities associated with the fall of the Roman Empire (Roberge 2010: 417). The latter, although, not exactly pertaining to the geopolitical reality of the late eighth and the ninth centuries, does resonate with the early Scandinavian talent for exploiting the shifting balance of power in the not-yet-united England (Roesdahl 1991: 189) and for recognising the dangers of impending Danish-Frankish confrontation[23] during the first decades of the ninth century (Geipel 1971: 36, Fisher 1973: 208-11). Apart from the causes identified by Roberge, there were a few more local ones, as highlighted by Roesdahl (1991). One of the forces driving the Scandinavians out of their native land could be poverty, which quite likely complemented the vicious circle along with overpopulation and competition for resources mentioned earlier. In addition, there were political issues to confront within the Scandinavian kingdoms, which either produced exiles[24] or gave rise to expeditions to gain means for maintaining kings' power. Next, Roesdahl speaks of the need to re-establish or to confirm the earlier connections. The diplomatic as well as trade relations which Scandinavia had with Britain prior to the Viking Age, mentioned earlier, would clearly fall under this heading. Finally, the incursions might have been stimulated by "a spirit of adventure, self-confidence and fatalistic attitude engendered by social conditions in Scandinavia" (Roesdahl 1991: 187-190). The development of "seaworthy and manoeuvrable" (Fisher 1973: 212) ships that could take a team of three dozen men certainly allowed Scandinavian fishermen and traders to become Viking pirates.

While Norwegian Vikings conducted the first lightning attacks on the English shores, the subsequent raids between 835 and 865 were carried out by the Danes, with large armies at their disposal and with settlement in mind (Logan 1983: 141, Loyn 1994: 39). Initially, the English were rather successful in withstanding the Viking raids. Their luck withered when the Danes decided to winter in England (from 850 to 851) (Loyn 1994: 38). From that moment onwards, serious Scandinavian incursions took place, culminating in the arrival of *mycel hœþen here*[25] in East Anglia in 865 (Roesdahl 1991: 234). The Vikings managed to take control over most of northeast England within the next five years and their eyes were soon directed towards Wessex (Fisher 1973: 222). This West-Saxon kingdom, however, proved to be a more formidable adversary than expected. During the reign of Ethelred and, subsequently, his younger brother, Alfred, the Vikings became aware of the limits of their determination and hunger for expansion. The decisive battle was waged at Edington (878), where

[23] Regular depopulation of eastern Holstein by Charlemagne brought Franks to the southern borders of Denmark (Geipel 1971: 36). Not only did Franks threaten the stability of Danish outskirts, they could influence, if not destroy, the well established Danish oversight over two important trading routes, which boosted the country's prosperity (Fisher 1973: 208-11).
[24] As a form of punishment for the unruly or the inconvenient.
[25] Great heathen army, as described in the Anglo-Saxon Chronicle (Geipel 1971: 40). 'Great' meaning two/three thousand men (Roesdahl 1991: 234).

the *here* was soundly defeated by the English. The Viking leader, Guthrum, was forced to sign up a treaty with Alfred at Wedmore. The freshly baptised Danish leader vowed to keep the Scandinavian activities off the Wessex soil, restricted to the areas ending east of Watling Street and north of the Thames. These north-eastern territories, where Scandinavian laws and customs prevailed, came to be known as the Danelaw. The *here* managed to conquer three of the four Anglo-Saxon kingdoms, it had land to live on and to farm it (Roesdahl 1991: 237). Indeed, it is roughly at that time (876) that we can speak of the earliest Scandinavian settlement in England, when the Viking army got hold of lands in Yorkshire.

As for the eastern Midlands, the territory was colonised roughly around the same time, the end of the ninth century. In both cases, the majority of inhabitants were of Danish stock. Norwegians took north-western counties as well as parts of Yorkshire (Gordon 1981: 326, also Oppenheimer 2006: 390). They arrived later and many of these Scandinavian settlers came from Ireland (see below) (Geipel 1971: 39). Much of Scandinavian military tactics in England relied not exclusively on their naval skills but also on the extensive use of already existing systems of Roman roads throughout the country. Interestingly, Vikings used horses merely to move from one place to another. Their wars were waged on foot (Logan 1983: 151). The Scandinavian conquest of (large part of) England was not a lucky turn of events for the savage and half-witted pirates[26], which, to many contemporaries, was an immediate image of the Vikings. They constituted an intelligent, brave and technologically advanced military force (Collingwood 1993: 85).

The Viking 'tentacles' seem not to have touched much of Wales to the extent they did England, Scotland and Ireland. Of course, there are instances of Scandinavian interference along the Welsh coastline, quite likely as a stepping-stone before or between various landings in Ireland[27] (Pollington 1989: 102, Logan 1983: 42, Loyn 1994: 37). The latter, on the other hand, experienced turbulence relatively throughout the Viking Age, with the Celtic opposition, for the most part, skilfully crushed by the early Scandinavians. The defeat of the Irish was an easy task to accomplish in view of a disorganised, non-unified territory, often engaged in local disputes. The Irish shoreline, rich with churches and monasteries ready to plunder, was, without doubt, a cherry on the cake for White-pirates (Norwegians). Suffice to say, the Scandinavians found Ireland attractive enough to colonise it and to found cities, serving as centres of trade as well as vantage points for further, not-so-diplomatic exploits, viz. Limerick, Wicklow (*Wiking-law*) and, of course, Dublin (Collingwood 1993: 77). As for Scotland, since Norwegians established colonies very early on in the Orkney and Shetland islands, it seemed logical that their spheres of influence should extend subsequently southwards, reaching Sutherland, the land south of their original landings. The first attacks on the northeastern part of England (Lindisfarne, Jarrow) could be seen as a by-product of this colonisation (Logan 1983: 40).

---

[26] Not that the Vikings would invent the phenomenon, at any rate. Piracy was a common practice among the native inhabitants and their neighbours along the sea coasts (Colliwood 1993: 52, Miles 2006: 159).
[27] As well as other coasts along the Irish sea, including the Isle of Man, the Inner and Outer Hebrides; Anglesey and later Normandy (Pollington 1989: 103, Logan 1983: 42, Loyn 1994: 37, Geipel 1971: 39). Iceland was not colonised until 874 (Collingwood 1993: 56).

In the Scandinavian district of England, the Danelaw, warriors were joined by the settlers (Fisher 1973: 243) and within the area, as Loyn (1994) puts it, there was a "consolidation of agrarian settlement and retention of military organisation" (Loyn 1994: 44). Indeed, the Scandinavian armies regulated colonisation and the inflow of immigrants. East Anglia formed two units under Scandinavian supervision, one at Thetford, the other at Colchester. The north-eastern part of the Midlands (former eastern Mercia) had a few sub-districts of the Danelaw, with fortified headquarters established at Northampton, Cambridge, Tempsford/Bedford and Huntingdon. The section of the Five Boroughs (Leicester, Nottingham, Derby, Lincoln and Stamford) was reinforced for the same purpose at that time as well (Logan 1983: 156). A high degree of colonisation can be inferred from the place-name and personal-name evidence as well as linguistic evidence (Logan 1983: 167, Fisher 1973: 243). Figure 2.2a below shows that the boundaries of the Danelaw were respected:

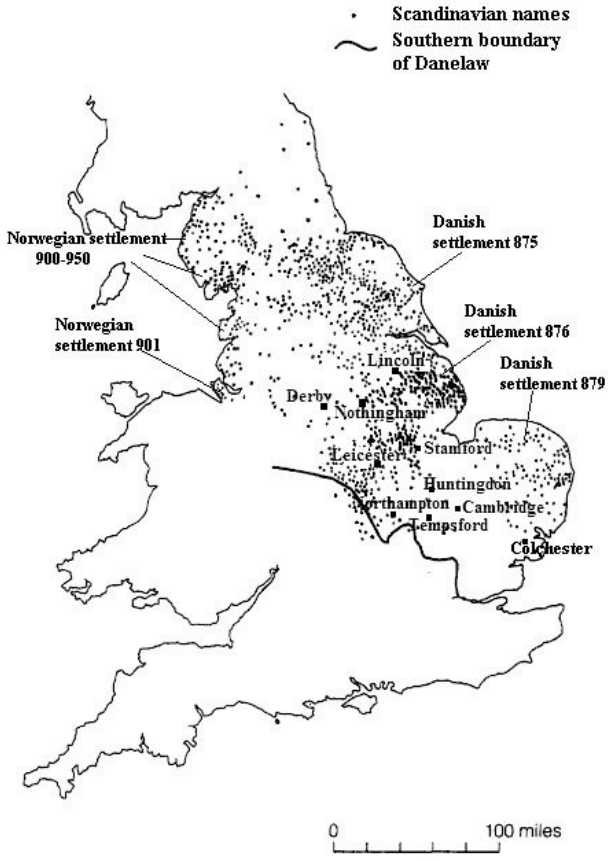

Figure 2.2a: Scandinavian Britain[28]

[28] Figure based on maps by Logan (1983: 169), Fisher (1973: 235) and Knowles (1997: 36).

Up north, the region was under the control of York, which was quickly growing into a potent Scandinavian fortified market. York was also instrumental in maintaining a dynamic settlement movement (Loyn 1994: 45), including the Vikings from Ireland, beginning with the tenth century (Fisher 1973: 238, 249). Admittedly, a reasonable level of tolerance on the part of the Danes was all the Irish-Scandinavians could expect. The Danish disdain for their overseas relatives was just as manifest as the hostility of Northumbrians for the Southerners (Fisher 1973: 248-9). In all, the area where laws were observed according to the Danish customs was efficiently organised, holding quite an individual society. One could easily distinguish a Scandinavian locality from an English one. The north-eastern shires were not divided into hundreds but into wapentakes; carucate was the unit of measurement of land instead of hide; the currency used by the (Anglo-)Scandinavian folk in the Danelaw was the ora, not the penny. Danelaw peasant population, too, enjoyed much more freedom than their southern English counterparts did (Fisher 1973: 243).

Meanwhile, south of the Danelaw, king Alfred dies and leaves two heirs to rule and protect England. His son Edward the Elder became the king of Wessex and his daughter Ethelfled ruled the English Mercia after the death of her husband. Both kept watchful eyes on the Danish matters and contributed to further unification of the country. From the end of the tenth century (980), however, the Isle experienced a second major wave of Viking attacks, which brought even more ferment and ended with the Scandinavian king Knútr (Cnut) on the English throne[29] from 1016 until 1035 (Logan 1983: 172, cf. Millward 1996: 69). The word 'Scandinavian' in Cnut's case, however, needs a proviso. Cnut constitutes a part of English history more than the Viking one. As a ruler, he divided the English territory into four earldoms along the lines of former kingdoms. Cnut himself watched over Wessex; Thorkell the Tall had East Anglia; Erik of Hlathir kept Northumbria. The infamous Eadric Streona was to oversee Mercia but this biggest traitor of Anglo-Saxon times was promptly put to death on Cnut's orders. Leofric took over as the Mercian earl. The *here* still resided in England but soon the political climate simmered down and there was no need for the great army to be any longer (Roesdahl 1991: 255).

Apart from reviving the English earldoms, Cnut set up a new English aristocracy. Surprisingly, only few Danish *jarls* were granted lands and the rest of the Scandinavian upper crust was sent back to Denmark. As a result, not many Scandinavian settlements were made, at least not to the extent comparable to the earlier colonisation (Gordon 1981: 326). Cnut's new nobility, therefore, consisted mainly of English-born collaborators or parvenus, who looked kindly on Danish overlordship and who replaced the English magnate families of the tenth century (Fisher 1973: 319-20, 338). In addition, Cnut made sure[30] that the mutual (Anglo-Scandinavian) responsibility, which constituted the basis for law and order in his country, extended from an

[29] The charter of 1027, ten years later, indicated a much wider extent of Cnut's sovereign power, making him the king of (all) England, Denmark, Norway and part of Sweden. Intriguingly enough, Cnut did not mention his authority over Orkneys and Shetlands nor did he declare his overlordship of the kings of Scotland and Dublin (Logan 1983: 177).

[30] By issuing the great code of law (1020-1023) (Fisher 1973: 332) drafted by Wulfstan of York (Fisher 1973: 335).

individual to a larger territory (Collingwood 1993: 160). In this sense, he was more English than Viking. Instead of the typically early Scandinavian manner of disregarding authority of the *jarls* and of shunning solidarity, except for battlefield scenarios (Pollington 1989: 20, Fisher 1973: 236 and 259, Crawford 2003: 67), Cnut reintroduced the Anglo-Saxon mode of binding communities, viz. the recognition of a kindred group and of collective responsibility (e.g. Fisher 1973: 122-3). What is more, he swore to promote Christianity, as another example of turning away from the Viking religion, and to protect the Church (Fisher 1973: 323). The fact that his high-mindedness made him honour the cult of the East Anglian king, Edmund, who was axed by the Danes [sic], as well as his assistance to bear the remains of the Archbishop of Canterbury, Alphege, who met the same fate, indeed, puts things into perspective (cf. Logan 1983: 178). It is important to pinpoint that attempts to establish friendly relations between the conquerors and the conquered was not exclusively on Cnut's agenda. From as early as the Alfredian times, there were various incentives to generate mutual tolerance and the Scandinavians were supposed to be treated as English subjects, not as enemies (Fisher 1973: 227, 249, 253, 302, 317, 331). Perhaps, the massacre of the Danes on St. Brice's Day (1002) ordered by king Æthelred, was an exception to the rule (Fisher 1973: 303)? After Cnut's death, his sons took over England. However, they proved to be much less skilled at the art of ruling the country than their father and, in the end, the throne was given back to a representative of Alfred's line, Edward the Confessor. Unfortunately, Edward died without an heir, commencing the ultimate contest for the sceptre of power.

The end of the Viking Age might be conventionally marked with the death of Harald Harðráði (Pollington 1989: 17), one of the contenders to the English throne, who lost to Harold Godwinson, Edward's brother in law, at Stamford Bridge (1066). The greatly disadvantaged[31] English forces met their fate a couple of weeks later at Hastings, fighting William the Bastard (re-dubbed Conqueror) and his army. The Normans, the last ones to invade Britain, although clearly of Scandinavian origin, represented already a distinct race (e.g. Barlow1966, 128), introduced a distinct line of political thought and spoke a distinct language (see section 2.4). It needs to be emphasised that the events leading to and including the fight at Hastings were not regarded as an ordinary struggle for the English throne between three princes, one of whom successfully concluded and secured his claim. It was a carefully planned venture especially on the part of William, with long-term goals in mind. Before heading for England, the Norman duke made sure he would get support for his military campaign from around (western) Europe, along with the acceptance of the pope. William, regardless of his claims being just or not, fought a holy war and everybody in the Continent agreed to his rights to seek divine vengeance (for the murder of Atheling) (Douglas 1966: 62-3). The diplomatic strategy and the approval that William received from European realms guaranteed the undisturbed reorganisation of England after the victory, the option which the other two princes certainly did not consider when launching their expeditions in 1066.

---

[31] Godwinson marched with his infantry from Stamford Bridge (Yorkshire) 190 miles south to fight the fresh Norman dismounted troops as well as the cavalry (e.g. Millward 1996: 69, Smail 1958: 134).

2.3 NORTH AND SOUTH: TWO ENGLANDS

The division into northern and southern England has a long-standing history. This divide has continued to linger until the present day (e.g. Millward 1996: 306), eminent at least in the language variation, at most in socio-economic and socio-political disparities (e.g. Upton 2006: 305-333, cf. Wales 2006, 33). Its beginnings can be found already during the Roman times, when the boundary between the civil and military districts was established (e.g. Blair 1977: 25). The boundary corresponded to the geographical split between the highland and lowland regions of Britain, with the natural frontier created by the river Humber. So well-founded was the division that the Anglo-Saxons, on their arrival a couple of centuries later, saw it fitting enough to respect it, up to the point of naming their kingdoms after these regional frames. The north-south partition, initially dictated by the natural features of the areas, plotted two distinct social realms, with differences building up as the subsequent waves of incomers peopled the Isle. Blair (1977) makes an apt observation on this development:

> *There are certain deep-rooted and long-lasting differences which distinguish the history of the northern peoples from the history of the southern kingdoms. The fundamental distinction is the one imposed by physical geography, (...), between a land difficult of access, of hills and moors from which indigenous elements could never be wholly rooted out, and an open, spreading plain upon which a new civilisation could be more easily imposed (Blair 1977: 26-7).*

It is upon these differences, too, that the foundations of the English language were laid. Surprisingly, the harsh, politically unruly north became the source of many structural linguistic innovations, including the aspects tackled in the current research. The multicultural character (e.g. Rowland 1993: 10) and the peculiar separatism of the North (Higham 1993: 194, Fisher 1973: 249) could serve as a catalyst for these linguistic changes.

During the days of the Heptarchy, the territory north of the Humber experienced a great deal of political instability. Initially the most powerful English realm, Northumbria went through disruptive currents – a sound defeat by the Picts (the Battle of Nechtanesmere in 685), subsequent civil wars and the ascent of Mercia. The ease with which the Vikings later took over the North must have been partially attributable to that political unrest (e.g. Loyn 1994: 24). The Church managed to thrive, nonetheless. What is more, the elevation of York into an archbishopric in 735 (see below) not only broke the unity of the English Church but also clearly confirmed the developing distinctions and isolationism of "the fifth part of England"[32] from the South. Even prior to the Viking attacks, the history of the North from the end of the seventh century, as noted by Fisher (1973), would diverge "steadily from the history of the southern kingdoms (...)" (Fisher 1973: 141 and 116). The contrast between the North and the South is striking and much of that divergence had to be built on hostil-

[32] Northumbria, as described in Egil's saga (Collingwood 1993: 49).

ity of one party towards the other, especially if the nuances of power were considered (Fisher 1973: 242).

Interestingly, the opposition was felt not only at lower social levels, which were not directly involved in the matters of the State, but also among the higher ranks (Fisher 1973: 271). Sources confirm that the West Saxon dominance was unwelcome up north and people above the Humber would always claim their strong independence of the southern views (Fisher 1973: 249). Similar disdain was expressed towards any potential candidate to the throne presented by the southern English (Fisher 1973: 317). There were attempts to appeal to the 'north way' by sending officials, whether of the Church or lay, but by and large, the southern kings kept themselves busy in their southern kingdoms and the northern magnates enjoyed their freedom far from the southern influence (Fisher 1973: 148 and 302). During the Danelaw times, too, the southern royals refrained from penetrating the northern lands, which enhanced the existing divisions (Fisher 1973: 275-6). In fact, as late as the thirteenth century onwards we can speak of regular exploits above the Humber by the kings (Fisher 1973: 347). Next, the North appeared distinct not only with respect to accepting the royal/administrative rule. Northumbria became a Christian realm quite early on, leaving other kingdoms behind, which made inter-state communication and cooperation difficult. The power of the North was feared, at that point, to such an extent that the (Christian) Welsh kings would ally with still-non-Christian Mercians to protect themselves from potential impact originating above the Humber (Fisher 1973: 108). The creation of two ecclesiastical provinces, on account of York becoming a separate archbishopric, moved the northern affairs even further away from the southern states, although the primary purpose behind the venture was merely to reform the organisation of the northern Church (Fisher 1973: 104). Christianity in the North, and soon throughout the rest of early medieval England, was a remarkable achievement and so potent that the raging Vikings could not extinguish it (Fisher 1973: 107). Remarkably, the ninth century northern Christian church was, in fact, instrumental in the conversion of the Viking Danes (Fisher 1973: 156). Further, the northern part of England was a place of refuge, whether for (politically) unwanted or persecuted. Anyone who was at odds with either the Scottish or the southern issues could find a safe haven in the North (Fisher 1973: 237 and 323, Rowland 1993: 15). These different individuals only added to the pot of already quite mixed northern population (Bugge 1921: 177, Fisher 1973: 249, Rowland 1994: 10 and 12). There were the British and the English, soon joined by the Scandinavians, all resided north of the Humber around the tenth century. Longstanding Northumbrian diplomatic relations with the Frankish kingdom, which initially served as a source of educational and artistic inspiration, subsequently extending into a full-fledged diplomatic cooperation, adequately enriched the cosmopolitan character of the North (Fisher 1973: 191). The result of the Northumbrian-Frankish contact was the evangelisation of Frisia. The northern missionaries moved there, as the Frankish influence expanded into that region (Story 1995: 76, Fisher 1973: 180). Their efforts must have proved fruitful, since sources mention a Frisian colony at York (Fisher 1972: 193), with flourishing educational exchanges (Fisher 1973: 188). Yet, of all the population groups residing in the North, the Scandinavians seem to have constituted the majority.

That the land north of the Humber was an established Scandinavian state at some point in early English history cannot be refuted. The list of 'festermen', men who gave pledges for Ælfric when he was elected to lead the see of York in 1023, confirms it, as does the Durham *Liber Vitæ*, the earliest section of which comes from the tenth century. Both sources are brimming with Scandinavian names, often more of Norwegian rather than Danish extraction. The former must have been long present in Yorkshire. Nevertheless, it is the Danes who were later more welcomed by Northumbrians, not the (Irish-)Norwegians (Collingwood 1993: 166). As mentioned in the previous section, the Scandinavian fractions were not always seeing each other eye to eye. It seems that the multiracial character of the northern territories spawned some complex internal issues, especially when they all had to be united under a single, English label (Fisher 1973: 249). While, the Norwegians formed many small isolated settlements there, the Danes annexed the English areas in a wholesale manner. Their settlements from the end of the ninth century onwards, not only (though essentially) in the North, had grave political as well as linguistic consequences for the English (Knowles 1997: 35, Wales 2006: 53). For once, the Scandinavian kingdom of York (and Dublin) had dismantled the old, Anglo-Saxon kingdom of Northumbria (Knowles 1997: 35). In due course, York would prove a substantial competition to its southern equivalents, Winchester and London, not only politically but also economically (Fisher 1973: 329, Stafford 2003: 38).

Interestingly, the territorial divisions in England set the North against the South much in the same way as they made the East Midlands run counter to the West Midlands. For the latter, too, contacts with the Scandinavian population determined the character of the eastern part of the Midlands area, especially with respect to the language. Fisher (1973) speaks of a linguistic frontier, which by the 11th century would break up the old Mercian kingdom into the English side (Oxfordshire) and the Scandinavian side (Leicestershire). The personal and place nomenclature serves as valid evidence in the case (Fisher 1973: 245). On the same note, Corrie (2006) points to the distinction between the eastern and western parts of the Midlands. She stresses the effects of two divergent histories, one with the Scandinavian input and the other without the Viking interference, which begin to emerge as early as during the first half of the Middle English period (Corrie 2006: 91). To complement the linguistic picture of medieval England, Knowles (1997) recognises a sharp contrast between the north and the east on one hand and the south and the west on the other after the Norman Conquest. The maps of LALME[33] confirm this regional variation (Knowles 1997: 44). The divisions, especially the one between the North and the South, had also a considerable impact on the linguistic situation (incl. language attitudes) around London, which, subsequently, had its role in the establishment of the national standard (see below and section 6.2.2).

[33] *A Linguistic Atlas of Late Mediaeval English* (Mclntosh et al. 1986).

2.4 AFTERMATH: SOCIO-LINGUISTIC SITUATION IN ENGLAND FOLLOWING THE INVASIONS

The coming of the Normans halted the continuous and arduous process of invasions and conquests. From now on, England could begin to exist as a country and military campaigns that subsequently predominated were largely those originating from her outwards, no longer the other way round (e.g. Smail 1958: 128). Internally, however, the English had to prepare themselves for serious restructuring and, unfortunately, for more ethnic cleansing, before the boundaries of the country and the foundations of power could be set for good.

After the victory at Hastings, William did not satisfy himself with taking the southern part of England alone. He had a clear vision of having the entire English territory at his feet, including the unruly North, willing to repeat his political success in Normandy (e.g. Douglas 1966: 65). He proceeded towards Yorkshire, destroying parts of the older Danelaw on the way. It is in the former kingdom of Northumbria that the greatest carnage was effected. William knew that drastic measures were the only option at hand, since the North was continuously in opposition (the rising of 1069-70), accepting the Scandinavian leadership with more enthusiasm, e.g. the welcomed landing of Ásbjörn's fleet in 1069 (Collingwood 1993: 174-5, Douglas 1966: 72). Astonishingly, the Harrying of the North, as it became known, and the aftermath, including the extreme cases of cannibalism of those who survived (e.g. Gillingham 2003: 214), did not eradicate the Scandinavian element from the north of England. The 1378 roll of freemen of York still shows surnames deriving from Norse nicknames. The area must have been repopulated, partly by immigration from Cumbria and Westmoreland, partly by the return of refugees from Scotland (Collingwood 1993: 177-8). Nevertheless, surviving the violent consequences of William's campaign was not the only trial of the remaining Anglo-Scandinavian population. Beginning from the twelfth century, the Anglian-Viking genes were further 'diluted' by the immigration from foreign countries, markedly as a result of development of various industries and the growing need for skilled workforce. There were not only Normans coming to the North but also the Flemish. During the fourteenth century, the northern society had been further enriched by the arrival of people from Scotland, Ireland, southern parts of England and, eventually, France (Colingwood 1993: 179). By the end of the medieval period in England, the North was a multilingual pack of nations.

The turbulent changes effected by William I were by no means confined to the North. The destruction, an immediate as well as expected consequence of the arrival of Normans, was comprehensive, reaching the majority of lands in England. Vast areas were wasted, many towns ruined, and soon, hunger and diseases finished the 'job'. Local rebellions were stirred up but no wins were noted on that account (Barlow 1966: 129). In fact, the uprisings brought even more damage, with the notable instance of the Revolt of the Earls in 1075, which foreshadowed the imminent collapse of the Old English nobility. The bloodshed of Fulford, Stamford Bridge and Hastings, followed by the uncompromised treatment by the Normans and, eventually, the substitution of the English nobles with the new aristocracy from overseas, crushed the English upper class. It has been reported that by 1086 as little as eight per cent of the English

lands belonged to what was left of the Old English elite (Douglas 1966: 65). If there was still English society after the Norman 'reorganisation', the English nobility ceased to be an essential part of it. It is not an understatement to claim that the new regime in England was introduced predominantly by force (e.g. Gillingham 2003: 215). After the ruin came the construction, of new Norman towers and castles, of French and German stylised cathedrals and churches, by utilising the southern (English) wealth, which now belonged to the new rulers. The ventures were carried out, too, mostly in the South. The North was not included in the Norman plans of rebuilding of the religious sites and it was certainly less considered when the project of royal castles and fortresses in the country was drawn (Gillingham 2003: 216-17).

Linguistically, the situation must have been quite peculiar at that time too, inviting the familiar divisions. In the South and East, the ruling elite spoke French while the minions used English. Up north (and west), however, where Norman claws did not hold so tight, both rulers and subjects still spoke the same language, viz. English (Gillingham 2003: 216-17). That being said, the distinction into the higher French and the lower English classes in the South would definitely not exclude cases of bilingualism (Knowles 1997: 47 and 50). In addition, if English was not appreciated it was more the sign of the times and of politics rather than a blatant act of repression. The original language of the Normans, Old Norse, faced a similar fate sometime in the past. In writing, some orthodox views on Normans destroying the English language seem not to be accurate either. Recent studies show that the French masters valued English writings, especially those referring to law (Knowles 1997: 48). Admittedly, while Latin was used as a language of documents, French was, as Knowles puts it, "the main auxiliary written language" in England until the fourteenth century (Knowles 1997: 47). What is more, from the mid 1250s, French language also began to take over the realm of court and of official records (Knowles 1997: 49). Yet, despite the inflow of new French-language written material during the first half of the Middle English period, the old Anglo-Saxon literary tradition still flourished within almost one hundred years after the Conquest. The bastion of earlier literary work, located in the southwestern part of the Midlands, continued to produce Old English texts at that time (Corrie 2006: 109). What eventually 'killed' the Anglo-Saxon style was the hand of the copyist. Old English ceased to operate because half way into the twelfth century, with the new linguistic climate building up, virtually nobody knew how to write it (Knowles 1997: 48).

The first signs of the rebirth of the English, both culturally and linguistically, should be sought during the beginning of the thirteenth century, with the loss of Normandy in 1204 and the Baron's Revolt of 1258 (Knowles 1997: 50). However, the main impulse towards the regaining of England, without doubt, was the Hundred Year's War[34] (1337), which left English and French as clearly two separate nations. As the war was waged, the English language steadily began to reclaim the roles formerly assumed by French. The piece of evidence showing that the French language was in

[34] With the Black Death of 1349 further decimating the society and with the Peasants' Revolt of 1381 providing another point of social turbulence (Horobin and Smith 2002: 28).

recession is Higden's *Polychronicon* from 1352[35], where he claims that the English language deteriorated through contacts with first the Scandinavians and later with the Normans: "(...), noþeles by comyxtioun and mellynge firste wiþ Danes and afterward wiþ Normans, in meny þe countray longage is apayred, (...)". He further maintains that it is the original mother tongue which should be used as the language of instruction as well as the means of communication among the elite. According to Higden, the fact that one is learning French at school does not mean they are using it for everyday purposes (Knowles 1997: 51). Soon after, English began to be heard at the royal court, with Henry IV (1399-1413) being the first king after the Norman Conquest whose native language was English. The language started to be used also during the legal cases, although the court's proceedings continued to be recorded in Latin (until the eighteenth century). The Lollard movement of the end of the fourteenth century became instrumental, among others, in bringing the native language issues to the foreground (Knowles 1997: 64). As English was being generally adopted in writing, more and more texts were made available to the public (Millward 1996: 184). This process also included the lower ranks of society. It is very likely that the increase in the availability of English written material induced the scribes to write and/or copy in their own local vernacular. It is in marked contrast with the literary tradition of the period up until around 1300, when texts would most likely be copied *literatim*, without any major amendments or embellishments on the part of a scribe (Corrie 2006: 101-2).

The linguistic renaissance coincided with socio-economic development. The fourteenth century witnessed the emergence of both a new group in English society, the middle class, as well as new plains where progress could thrive on, the cities. The new class comprised mainly merchants, manufacturers and traders. They operated in urban centres rather than in rural areas. London, the country's new capital, became one these centres[36]. It was a place of immense growth and powerful influence for England and its people, within as well as outside the country (e.g. Horobin and Smith 2002: 29). The merchants saw opportunities for cooperation not only between various regions in the Isle but also with the foreign lands. As the trade connections flourished, the use of English became vital in all aspects of merchants' work, including keeping records of guilds (from 1380s onwards) (Knowles 1997: 52). It was a vibrant time when French, English and Latin would occur together, especially in writing.

[35] Translated into English by John of Trevisa in 1387 (Knowles 1997: 51). The piece represents the specimen from the southern Middle English dialect.
[36] Along with York, Oxford, Norwich, Glocester (Horobin and Smith 2002: 29).

3 Word Order Shift: Causes, Mechanisms and Factors behind the Change

3.1 DETERMINANTS OF WORD ORDER: FROM FLEXIBLE TO FIXED

By definition, when dealing with word order, we speak of "the linear order in which words are arranged in sentences" (Song 2011b: 254). If linear fashion is not observed, a given sentence is rendered ungrammatical. The example in (1) below (from Dryer 2013a) presents the ordering of words in English that is considered basic, with three major constituents, subject (S), verb (V), object (O):

(1) The dog chased the cat.
S (NP) V O (NP)

SVO in English is considered rigid. The roles of the noun phrases (NP) are defined by their particular positioning relative to the verb. SVO is also a dominant order in the English language. It is the most frequently[37] represented type, and any departures from it are either restricted for pragmatic reasons or are simply not grammatically viable (cf. Dryer 2013a). Looking at the example above, one would identify the basic word order within a stylistically neutral, independent and indicative clause as the one involving "full noun phrase participants, where the subject is definite, agentive [A] (...), the object is a definite, semantic patient [P], and the verb represents an action, not a state or an event" (Siewierska 1988a: 8). Mithun (1992) sees the basic word order as "a primary characteristic from which other features of the language can be predicted" (Mithun 1992: 15).

Pragmatic neutrality, transitivity and (textual) frequency aside, other criteria for determining the basic word order may also include formal markedness (e.g. Song 2011: 254). However, the 'basicness' is not so clearly identified in all languages of the world. With the so-called flexible order (free order) languages, the criteria presented above may not suffice. What is more, it has been frequently emphasised that the term 'free' is inadequate in word order type nomenclature because there are usually pragmatic and/or syntactic factors involved that determine a language's linear order. Languages with flexible orders can be further classified into those that have a dominant order (Russian) or those that lack one (German, Dutch) (Dryer 2013a). Those that lack one can exhibit high flexibility, where all constituent permutations will be possible and operational, e.g. Nunggubuyu, aboriginal language spoken in northern Australia (Heath 1984: 507-513; 1986). There are also types which do not have a

[37] WALS uses the term dominant order instead of basic order, focusing on the frequency of use, as reflected in texts (Dryer 2013b). For cases where basic order may not equal with a dominant one, see Abbott (1991: 25).

dominant order simply because the subject or object show flexibility with respect to the verb, e.g. Syrian or Arabic (Cowell 1964: 407, 411). Moreover, Mithun (1992) shows that the criteria for detecting the basic word order actually do not apply to instances of flexible word order. She provides examples of pragmatically based languages such as Cayuga, Ngandi, Coos, where word order(s) is (are) determined on "the relative 'newsworthiness' of the constituents to the discourse" (Mithun 1992: 39) - on bringing in a new topic, new information, or indicating contrast.

The fact that some languages do not exhibit a clear preference for one particular word order made scholars explore this inconsistency. Early on Lehmann (1978b) proposed that languages which showed the properties of both OV and VO were regarded as departures from ideal OV or VO states and were actually in the process of changing to either of the two. The changes, he claimed, could occur due to external pressures such as language contact or a language's internal development (1978b: 32-7, also Andersen 1983: 19-20, 21). While scholars have managed to convincingly show that Lehmann's 'inconsistent' languages constitute yet another type, drifting of languages towards consistency rather than inconsistency cannot be refuted (e.g. Song 2001: 304, Cristofaro 2011: 237, for exceptions and arguments on consistency see e.g. Deutscher 2000 on the case of Akkadian). What is more, even if Lehmann failed to acknowledge what Song (2001) describes as "the graduated distinction between more probable and less probable language states" (Song 2001: 310), he was not wrong by pointing to both internal and external pressures being instrumental in syntactic changes, word order shifts included (e.g. Hickey 2002 on the Celtic-English contacts).

Word order 'freedom' may involve uncertainties around comprehension. These uncertainties can be resolved through coding mechanisms, especially overt case marking. Presence of case is crucial with respect to providing information on the syntactic function of arguments. It is also involved with forming expectations on the argument make-up within the predicate of the sentence, as case makes an integral part of certain constructions or lexical items (Lamers and de Swart 2012: 4-5). Indeed, scholars often point to patterning between word order flexibility and the presence of case marking (e.g. McFadden 2003: 295, 300, see next section). Still, for many languages a straightforward correspondence between the two phenomena cannot be ascertained. The impossibility of such correspondence needs to be at least recognised in view of the lack of equivalence of (functions of) cases across different languages (Blake 2004: 155, also McFadden 2003: 299). Furthermore, there are more factors to take into account when identifying grammatical relations. There are also other scenarios to consider when these relations become underspecified. Some of them may not necessarily be realised through word order changes. Conversely, explanations on word order shifts cannot always be explained by a single reason or conditioned by one particular driving force (e.g. Seoane 2006: 362). It is worth mentioning, though, that word order is often regarded as an obvious alternative to case marking in distinguishing subjects from objects (e.g. Blake 2004: 14). What is more, when cases erode, some changes in word order may occur as a prophylactic measure[38] to decrease the

[38] The compensatory nature of implementing (subject-object contrasting) word order should be distinguished from the proactive application of word order strategy when case is persistently employed in a language (e.g. De Vogelaer 2007: 176).

ambiguities created by the lack of an overt case marker (De Vogelaer 2007: 169), as suggested by diachronic evidence, e.g. from early English (Bean 1983: 139, cf. Fischer 2010: 67). Finally, if word order shift ensues to replace case, especially to differentiate between subjects and objects, the (strict) SVO element arrangement seems the most favoured option (e.g. Hawkins 1986: 48-9, Blake 2004: 158-9, cf. Hickey 2002: 270, De Vogelaer 2007: 174).

It was pointed out earlier that although an interaction is to be expected between the word order flexibility and the presence of case, other phenomena need to be considered when interpreting word order variation. One of them is (verb) agreement, which along with case constitutes morphological means to differentiate subjects from objects (cf. Deutscher 2000: 60). It is interesting to observe how the three parameters for defining grammatical relations, namely (SVO) word order, case and agreement pattern within the Germanic language family, as presented by De Vogelaer (2007: 179, table 7). German, Yiddish and Insular Scandinavian appear as the most conservative among the Germanic languages, i.e. they have moved away little from the original Proto-Germanic configuration. This configuration encompassed the presence of SOV word order, case and verb agreement. All three languages have case productive on NPs and there is verb agreement at hand. The only difference between them is how SVO patterns in clause types. In German, SVO is found only in the main clauses, whereas in Yiddish and Insular Scandinavian, SVO is operational in both main and subordinate clauses. On the opposite end, there are Afrikaans, Continental Scandinavian and English. They exhibit SVO in both clause types, with verb agreement either absent entirely or almost non-existent in some/most varieties and with remnants of case, usually on pronouns. Between the conservative and innovative Germanic languages are Dutch and Frisian. For these two languages, SVO is present only in main clauses. In addition, while the verb agreement is operational, case (mostly) is not.

De Vogelaer rightly notes that all languages in the Germanic family utilise at least one of the three means to define grammatical relations (2007: 180). It also follows clearly from this description that case as a morphological device does not stand on its own but is often employed along with verb agreement. Moreover, instances such as Dutch and Frisian appear to act as counterexamples to claims of correlation between case marking and word order flexibility. Verb agreement looks like a sufficient marker of grammatical relations in these languages (cf. McFadden 2003: 304-5, but see below). It should be emphasised that Dutch and Frisian together with English are the only languages in the Germanic family which use a single mechanism to distinguish subjects from objects. Yet, English, as De Vogelaer (2007: 177) points out, has a strict SVO, accompanied by phenomena not present in other Germanic languages, such as raising, extraction, and NP-deletion. Interestingly, even if the same rigid syntax cannot be distinguished in English and Dutch, some restrictions on Dutch word order have to be considered, especially when compared to that of German[39], which still uses morphological case (cf. De Vogelaer 2007: 176). In this respect, one can still place

[39] Dutch (and Frisian), unlike English, still use V2 in main clauses and in the subordinate environment, they still keep SOV. These characteristics keep the two languages more on the conservative (German) rather than innovative side (English) (e.g. De Vogelaer 2007: 171-2).

(rigid) SVO against one of the two or both morphological means of identifying grammatical relations.

Other aspects have to be taken into account as well, when looking for causality and correlations between instances of grammatical marking within the Germanic family. De Vogelaer includes, among these, the 'pro-active' versus compensatory means of introducing SVO word order (Insular Scandinavian vs. English and Afrikaans respectively). He also stresses the fact that rigidness of word order and the loss of case do not have to occur at the same time (e.g. Dutch vs. Icelandic, Yiddish)[40]. Another confounding factor is how the verb-second constraint is realised within the Germanic family. De Vogelaer mentions the ongoing dispute whether to regard embedded clauses in the Continental Scandinavian as SVO-clauses or clauses with V2 (2007: 176 f11). His observation echoes the remark made by Mitchener (2005) with respect to the syntactic situation in Middle English. In declarative clauses with topicalised subjects, SVO grammar with V2 could create the same surface word order as an SVO without the verb-second constraint (Mitchener 2005: 11). Accounting for these aspects does not make generalisations easily valid (see next section for details). They also cannot completely rule out case as a vital determinant in the framework of grammatical relations.

It is interesting to see English, the object of the present study, assigned to the innovative group of the Germanic family[41], along with the Continental group of Scandinavian languages and Afrikaans. The former finds a common trait with English, among others, in the lack of case effects on word order. Primus (1998) observes that loss of case marking in all these languages altered both overt and syntactic case distinctions. The distinction between nominative and objective is now structurally expressed. As for Afrikaans, it evolved in a similar manner. Typically for a contact language (Matras 2009: 281-2, Thomason 2001: 167), it has no case and agreement (also Romaine 2002: 590), with SVO order present overall as a compensating and not pro-active strategy (e.g. De Vogelaer 2007: 179). It would be tempting to ascribe the relationship between the lack of overt morphology and the presence of (compensating?) rigid SVO in English to contact-induced pressures which the language underwent throughout its early stages of development (including those with the early Scandinavian population). Indeed, various scholars over the years have pointed to far reaching contact-induced outcomes on the language at this stage, some even referring to Middle English as a creole (Bailey and Maroldt 1977, Poussa 1982, Thomason and Kaufman 1988, cf. McWhorter 2002, Roberge 2010: 421, Trotter 2012). The attractiveness of such claims could be backed by the fact that some of the contact scenarios proposed for English involved a substantial loss of inflectional morphology (Fisiak 1977, Poussa 1982, Milroy 1984, Danchev 1988 (1991), Ruiz Moneva 1997. Still, numerous studies have shown that both internal and external factors have to be con-

[40] One may precede or supersede the other, e.g. Dutch – morphological case disappears, more rigid word order continuously ensures (De Vogelaer 2007: 176); Icelandic (and Yiddish) – case still productive in the language, with pro-active SVO recently developed (cf. Hróarsdóttir 2000: 259).

[41] An intriguing proposal has been recently put forward by Emonds and Faarlund (2014), who claim that (Middle and) Modern English have not developed from Old English. Instead, they propose that Middle English is a form of "Anglicised" Norse, whereby Norse supplanted Old English (2014: 11, 154, 156). This would place English closer to the North Germanic branch rather than the West Germanic one.

sidered here. The following sections further explore the nature of morphology and word order interaction as well as provide a broader overview on internal and external motivations for language change.

3.2 MORPHOLOGICAL SIMPLIFICATION AND CHANGES IN WORD ORDER

The study of relationship between (rich) morphology and a word order type is a long-standing one. Greenberg's Universals of Grammar (1963a) contains one of the earliest significant attempts to establish a connection between word order and the type of case marking. His Universal number 41 stated that "if in a language the verb follows both the nominal subject and nominal object as the dominant order, the language almost always has a case system" (Greenberg 1963b: 96). Greenberg's study of thirty languages of the world, indeed, spoke in favour of a correlation between the presence of case marking and SOV. Numerous studies to examine the relationship have been carried out henceforth. The ideas presented by Lehmann (1978) already mentioned in section 3.1, are but one of the relevant instances. Next, there was the study conducted by Mallinson and Blake (1981) where as many as a hundred languages were examined. They found, among others, a correlation between the preference for SOV and the presence of case marking on one hand and between the preference for SVO and the absence of case inflections on the other (Mallinson and Blake 1981: 179)[42]. Similarly, Siewierska and Bakker (1996) explored how case along with verb agreement pattern with word order types. They had a large database at hand, with as many as 237 languages examined. Of the correlations investigated, they found one between SVO and the absence of verb agreement. This correlation is accounted for via the assumption that, much like case and SVO order, agreement enables the identification of grammatical relations (cf. De Vogelaer 2007: 173). This characteristic transpired already in section 3.1, when the occurrence and co-patterning of these three coding strategies was presented for languages within the Germanic family. Siewierska and Bakker's recent study (2008) continues to explore the co-occurrence and functions of these strategies. This time, their sample included as many as 417 different languages. They used the tripartite typology of verb-final, verb-medial and verb-initial languages, with APV and PAV belonging to the verb-final group. They recorded, among others, "dispreference" (2008: 296) for case marking in verb-medial languages and the preference for case marking in verb-final languages. The former can be explained to be due to an overlap in function of the two means of syntactic encoding. Once Agent (A) and Patient (P) are placed on the opposite sides of the verb, the marking of grammatical relations through case can be regarded as "superfluous" (Siewierska and Bakker 2008: 296). Similarly, Primus (1998) comments on the presence of case within the configuration where verb intervenes between A and P, rendering the former irrelevant (1998: 447).

[42] Observations on the presence/absence of case marking for VSO type could not be made. There were too few VSO languages in the sample.

A recent study by Bentz and Christiansen (2013) also explores, among others, the validity of Greenberg's Universal 41 in WALS. Here, the universal is checked by crossing the features (49A) "Number of Cases" and "Order of Subject, Object and Verb" (2013: 58). Their investigation confirms the link between SOV and presence of overt case as well as "the dissociation of SVO languages with case marking", which both "are general trends that hold for a wide sample of languages across the world" (Bentz and Christiansen 2013: 58). Eighty per cent of SOV language have two or more cases at hand, while seventy two per cent of SVO languages types display no morphological case marking.

The existence of a relationship between a rich morphology and a particular word order type naturally extends to discussions around how much shifts in inflectional paradigms influence word order change. Some observations on that have already been made in the previous section. The hypothesis pursued in this study is that the loss of inflectional morphology facilitates and accelerates word order change. Indeed, the fact that syntactic change is somehow connected to morphological change taking place earlier is a well-recognised and studied idea. It is thought to be connected to the complex interaction between the loss of information which follows morphological simplification and the increase of information which occurs when more rigid word order rules apply (cf. Bentz and Christiansen 2013: 47). It has to be noted, too, that theorising around change in morphology cannot exclude the possibility of an increase in morphological complexity as well. Some scholars would see this as proof of a stronger link between morphological and categorial changes (e.g. Lightfoot 2006a: 35). Another, more absolute, claim to the effect that loss of inflections drives all the changes in linearization can be easily dismissed. Decrease in richness of inflectional morphology alone cannot be responsible for, as Li (1977) puts it, "the most drastic and complex category of syntactic changes" (1977: xii). Word order change stirs the fundamentals of the syntactic framework of a language. Such a change involves a myriad of coexisting or consecutive structural modifications which operate in an organised fashion so that a language shifts from one type to another. The results of such a change are not limited to a simple reshuffle of the order of subject, object and verb in a given sentence. It has to be remembered, too, that the full realisation of a word order change occurs over a very long period of time (Li 1977: *idem*). As an illustration of this, there were various processes involved in the Old to Early Modern English word order change. The majority of them were syntactic in nature. Apart from the morphological simplification, there was the grammaticalisation of afterthought material and the issues around constituent weight, to mention but a few (Seoane 2006: 360). Finally, the third scenario, which presents morphological simplification as having no impact on word order shift, seems somewhat unreasonable at a first glance. The very essence of inflectional morphology is "to symbolise grammatical categories in the word", as remarked by Wurzel (1989: 179) and any significant alteration in its status is bound to have repercussions in syntax. However, despite the recognised relationship between these phenomena, studies show that the two can also change independently. Indeed, Kroch (2003) remarks that syntactic change does not always have to be preceded by changes in morphology (2003: 399). It all boils down to individual languages and their idiosyncratic historical paths.

To establish how much inflectional morphology conditions word order change means, to a large extent, trying to find an explanation to the correlation between a rich case marking system and word order freedom. One can lean on the perspective offered by the principles of synchronic grammar, which concern derivation and representation, with particular premises expressed formally. These principles, regulating word order, refer to morphological case, whereby rich case marking licenses given word orders that are not viable when this marking is not overtly present. Yet, this kind of perspective is not the only one available. It is also possible to try and work out the nature of the correlation in terms of use, acquisition and change. Thus, word order variation is easier to comprehend, acquire and maintain over time in a language where grammatical relations are marked with morphological case. These two distinct approaches to the case-marking/word order freedom correlation (CWC) are examined by McFadden (2003). He refers to a few synchronic-grammatical explanations offered by Roberts (1997), Kiparsky (1997), Neeleman and Weerman (1999). Three aspects of the CWC are presented and analysed. The first one states that the CWC must be considered in terms of the syntax-morphology interface. In other words, within the synchronic grammar, should morphological case influence word order the syntax has to at least partially depend on morphology (McFadden 2003: 296). The second characteristic covers the idea that the CWC entails optionality (McFadden 2003: 298). The third aspect explores the degree of richness of case marking and how strong it has to be to allow word order freedom (McFadden 2003: 299).

The main assumption behind McFadden's two-fold analysis is that attempts to explain the correlation between case marking and word order flexibility within the framework of synchronic grammar can be questionable. One reason for it is that, unlike scenarios on the CWC thorough principles of acquisition and use, synchronic-grammatical accounts depend on, what McFadden calls "a synchronic causal connection from morphological case to syntax" (McFadden 2003: 298). Some of these synchronic accounts also find the optionality which word order flexibility involves difficult to account for (e.g. Roberts 1997). Otherwise, even if these synchronic proposals aim at generating an optionality effect (e.g. Kiparsky 1997, Neeleman and Weerman 1999), the theoretical stages employed to derive that optionality are, as McFadden argues, problematic[43]. On the other hand, the use-based approaches to the CWC can easily explain optionality and how speakers deal with it (McFadden 2003: 299). Finally, synchronic scenarios require a formal definition for a 'rich' morphology that applies universally, whereas a use- and acquisition-based approach does not need to incorporate precise definitions to establish how rich morphology has to be to license word order freedom (McFadden 2003: 300). Instead, it relies on the ease with which a speaker-hearer uses and understands utterances. The notions such as 'sufficiently distinct' and 'unambiguous' are made available at his point.

Before providing his scenario of word order change, McFadden concludes on the basis of empirical data that the CWC is not a (one-way) implication but a tendency (McFadden 2003: 304). This tendency can be seen as resulting from principles of lan-

[43] Some aspects of these synchronic accounts have been abandoned due to intricacy they carry into grammar or due to the fact that they do not conform to the fundamental assumptions around morphology (McFadden 2003: 299).

guage use, acquisition and change. Thus, McFadden sees the progression of word order change in a language with a high degree of word order freedom, such as that occurring in early English, as a 'cruel fate' of the more marked word orders, once the case marking was lost. With the ensuing morphological attrition, these orders would become increasingly difficult to interpret correctly. As children would find it hard to acquire the processes which generate these orders, the increasing avoidance of the marked orders would be employed to guarantee being understood. As a result, the marked orders would become more marked. In order to be interpreted clearly and to avoid ambiguities, the marked orders would increasingly need stronger pragmatic and intonational motivation. Finally, children would not acquire the word order freedom because there would not be enough evidence for the marked orders in the primary linguistic data (McFadden 2003: 304-5). McFadden concedes, quite rightly, that other factors have to be considered as well when explaining variation of word order. Among them are, expectedly, agreement and intonation. They enable existence of certain word order flexibility even in the absence of case marking. At that point, the intermediary position of Dutch within the Germanic language family (see the discussion in the previous section), with respect to particular type of grammatical marking, becomes justified.

That the loss of (case) inflections makes certain orders (less or) more dominant is also remarked on by Allen (2006). At the same time, she observed, based on data from early English, that the growing rigidness of word order could not be conditioned exclusively by deflexion. This takes us back to one of the three statements about the impact of the loss of morphology on word order change. At the same time, Allen concedes that, for this particular language instance, the two developments might have worked in tandem and that such a scenario agrees with current Case licensing theory (Allen 2006: 220).

Hróarsdóttir (2009), too, explores the correlation between word order shift and absence/presence of a strong morphology, focusing on English and Icelandic. She claims that the strength of the link between a rich morphology and OV order, in particular, has been exaggerated and that it can only proceed one way. For languages that have shed their rich morphology, the loss of overt OV can indeed take place, e.g. English. However, the course of changes in reverse direction is not workable. In this respect, some languages may exhibit overt OV regardless of the presence of (a weak) morphology. Furthermore, languages may discard OV whether any changes in morphology actually occur alongside or not, e.g. Icelandic (Hróarsdóttir 2009: 67-8). The fact that the loss of inflection may prompt a grammar change is connected to language learners' paying attention to "a universal cue for the positive value of the OV/VO parameter" and to how different languages express this cue (Hróarsdóttir 2009: 67). For languages such as English, the cue may have been realised through morphology, making the grammatical shift feasible. By contrast, in Icelandic, the cue for OV was expressed through information structure (Hróarsdóttir 2009: 88-7). Interestingly, for both languages, external effects contributed to fewer expressions of the (universal) cue, thus a change in grammar occurred, though at different times and in a different manner.

McFadden's account on the CWC as well as Hróarsdóttir's comparison of English and Icelandic presented above have also been brought to the fore because they stress

the significance of acquisition phenomena in the process of morphosyntactic change. Hróarsdóttir (2009) specifically refers to Lightfoot's (1999) cue-based approach[44], whereby a cue, i.e. a structure, is created in a child's I-language grammar after being exposed to the primary linguistic data. A language change occurs when the use of a feature (in her case OV word order) decreases until the frequency of occurrence falls below a certain threshold to be useful as a cue (Hróarsdóttir 2009: 80). The mentioned study by Bentz and Christiansen (2013) is another example of how acquisition matters for language changes. They investigate what they refer to as the trade-off between nominal case marking and fixed word orders (2013: 45). The focus of their study is Germanic and Romance languages. The study highlights the production constraints as a source of linguistic change. Bentz and Christiansen claim that the impact of L2 learning constraints is a relevant, though not exclusive, factor for the receding case paradigms. These constraints revolve around language production which non-native speakers have to confront. The problem at issue is the difficulty in recognising the correct nominal affixes that learners of Middle English (among others) would have experienced when producing even simple transitive sentences (Bentz and Christiansen 2013: 52). The stronger "trade-off" between SOV with case marking and SVO without the overt cases was created as an effect of confusion between formerly distinctive case declensions. The non-native speakers "recruited" (2013: 55-6) for the change to take effect were the early Scandinavians and, later, Normans arriving and settling in the British Isles.

3.3 MOTIVATION FOR CHANGES IN MORPHOSYNTAX

The distinction between two types of explanations for the relationship between word order freedom and morphological case marking made by McFadden (2003) outlined in the previous section, touches upon the existence of a dichotomy between formal and functional explanations of linguistic phenomena in general. In the former, as described by Newmeyer (2003), the key role is attributed to principles that govern the organisation of *grammars* (2003: 18). The latter deals predominantly with properties of languages *users*. These properties refer to the users' "interest in producing and comprehending language rapidly, to their states of consciousness, or to aspects of their behaviour" (2003:18). Thus, changes that are functionally motivated are often seen as the ones that simplify, increase the efficiency of production, perception and acquisition (Campbell 1981: 172). By contrast, formal explanations move away from language users and tend to treat the language as a self-contained, autonomous entity (e.g. Heine 1994: 256). It is not surprising, therefore, that a formal-functional dichotomy is often paralleled by a distinction between internal (intra-systemic) and external (contact) explanations for language change (e.g. Heine 1994: 256). The two

[44] As remarked by Lightfoot and Westergaard (2007), the cue-based approach is a modification of the principles-and-parameters approach to language variation and acquisition. It differs from the latter approach in that "children do not evaluate grammars against sets of sentences"; they "are insensitive to the set of sentences generated by any grammar" (2007: 397). Essentials on the cue-based learning are presented in Lightfoot 2006 (2006b: 66-86). Comments on the Lightfoot's earlier assumptions under the framework of principles-and-parameters are found in sections 4.3.1.2 and 4.3.2.

often operate in tandem and the relevance of multiple causation cannot be denied (e.g. Thomason and Kaufman 1988: 57, Thomason 2001: 62, 91). Finally, Farrar and Jones (2002) stress the role of extra-linguistic (i.e. socio-political and economic) motivations in language change, how they interact with language internal and external pressures (Farrar and Jones 2002:1). The following sections outline the essentials on these pressures, as relevant especially for changes in morphosyntax. With English being the focus of the present study, most of the description concerning the internal as well as external pressures will be presented from the perspective of the developments within the Germanic language family.

3.3.1 Internal pressures

The progression of morphological changes in Germanic languages has been conditioned by two primary trends, simplification and grammaticalisation. Simplification entailed reductions in the number of categories realised by inflection. The change operated within two dimensions. On one hand, there was either an attrition or elimination of the categories themselves (e.g. nominal case, person, number inflection on verbs). On the other side of the spectrum, there was the supplanting of morphological expression with a syntactic one (e.g. the modal system) (cf. Faarlund 2001: 1717). This kind of change has been associated with drift, a term first coined by Sapir (1921), whereby "a language moves down time in a current of its own making" (Sapir 1921: 150). He listed three particular drifts as occurring in the course of changes within the English language. The first one was the levelling of the distinction between the subjective and objective cases (Sapir 1921: 163). The second was "the tendency to fixed position in the sentence, determined by the syntactic relation of the word" (1921: 166), which advanced the idea of the rigid SVO order. The third drift was that "towards the invariable word" (1921: 168). The first two drifts, according to Sapir, were essentially related. In his view, what starts as a small phonetic readjustment might, in due course, produce a major structural change (Sapir 1921: 174).

Understanding drift can be problematic from the point of teleology. While serving as "a pathway for change, a manner of change", drift does not constitute a cause (e.g. Harris 1981: 191). In other words, it explains how particular developments occurred but it does not provide answers to the question why they ever took place. Various scholars, according to Harris (1981), have often inadequately made a distinction between these two aspects of the change (Harris 1981: 188). More recently, Keller (1994), too, agrees that drift cannot be considered a reason why certain processes occur in a language. Still, he does not deny the existence of drift:

> *"[A] drift comes about when speakers are faced with certain ecological conditions, arising from (among other things) the respective state of their language and the problems inherent to the success of communicative enterprises, and react over and over again according to the same maxims" (Keller 1994: 109).*

Similarly, McWhorter (2002) considers the notion of drift more as a description than an explanation. To illustrate his logic, McWhorter questions the force of "general shift" and wonders why all Germanic languages except English managed to keep a gram-

matical gender distinction in the articles, having undergone the erosion of nominal inflection (McWhorter 2002: 229).

Although Sapir (1921) referred to English and to the language-specific phenomena, Lakoff (1972) along with several other scholars over the years have shown that these separate phenomena are, in fact, part of larger typological shifts. Lakoff sees the drift as "historical fluctuation between syntheticity and analyticity, acting as a sort of linguistic pendulum" (1972: 179). Thus, within the Germanic language family, drift entails shift of languages from more synthetic to more analytic types, with each of the languages occupying their own place on the syntheticity-analyticity scale. The diachronic aspect to this particular distribution, mentioned earlier in section 3.1, describes the early to middle stages of development of these languages that were more synthetic (also Faarlund 2001: 1708-9). Moreover, Lakoff regards drift as "a metacondition on the way in which the grammar of a language as whole will change" (1972: 178), an observation which, according to Hopper and Traugott (2003) separates the Sapirian drift from unidirectionality of grammaticalisation. While drift deals with regularisation of construction types in a language, unidirectionality has to do with changes that affect particular types of construction (Hopper and Traugott 2003: 100).

With grammaticalisation, we reach a second trend that affected the morphology of Germanic languages. Specific for this language group, it involved a phonological and semantic reduction, which has predominantly affected verbs. They have been reduced to auxiliaries semantically and in some cases to clitics phonologically (e.g. Modern English) (Faarlund 2001: 1717, Newmeyer 2003: 23). More broadly speaking, grammaticalisation is regarded as a complex change, a process which involves not one component but which proceeds across different components, with contributory changes occurring in phonology, morphology, syntax, and semantics (e.g. Fischer 2010: 13). The changes that grammaticalisation entails are said to follow the pathway of a cline. Hopper and Traugott (2003: 7) present it as follows:

content item > grammatical word > clitic > inflectional affix

This cline, therefore, represents a path from less to more grammatical (e.g. Bybee *et al.* 1994: 9-22). The unidirectionality within grammaticalisation is considered to be, in the words of Haspelmath (2004) "[t]he most important constraint on morphosyntactic change" (2004: 21). However, much like with any theoretical construct, there are a few caveats when accounting for grammaticalisation, as outlined by Faarlund (2008). Firstly, there are changes which proceed in the opposite direction from the one promoted on the cline (Faarlund 2008: 224). Secondly, explaining language change by grammaticalisation separates the language from speakers and learners, making it an independent object (Faarlund 2008: 224). Thirdly, grammaticalisation, according to Faarlund, has a poor explanatory value (2008: 226).

Meillet, who first coined the term grammaticalisation, proposed that the process of grammaticalisation includes word-order changes (1912 [1965: 148], similarly Kuryłowicz 1965 [1975], Givón 1979a). Many recent studies treat word order change as a result of grammaticalisation. Within this approach, there are proponents who see word order shift as a categorial change (e.g. Anderson 1986, Claudi 1994). On the other

hand, there are studies advocating word order change as something that is induced by changes in morphology, with grammaticalisation of e.g. periphrastic constructions into non-periphrastic ones involving transposition of constituents (e.g. Roberts and Roussou 2003, van Gelderen 2004). At the opposite end, there are linguists who see the two processes as separate, for example Hopper and Traugott (1993). They admit that changes in word order can have an extensive impact on the grammatical rules and texture of a language but as an example of reanalysis, these shifts should not be equated with grammaticalisation because they are not unidirectional (1993: 24). What is more, according to Hopper and Traugott, word-order shifts cannot be regarded as instances of grammaticalisation, as the wide array of reasons for word order change includes language contact, which is not a linguistic factor (Hopper and Traugott 1993: 63).

As regards grammaticalisation and language contact, these two phenomena can coexist. What is more, the former may be reinforced by the latter, for example, if a language exhibiting processes of grammaticalisation influences another language that does not have them (e.g. Wiemer and Bisang 2004: 12). The following section tackles contact-induced effects, including structural interference.

3.3.2 External influences

If one was to identify a pressure which could disturb "a current of [language's] own making", especially with respect to the pace of linguistic change, external influences would be the most immediate 'culprit'. Indeed, language contact, a notable instance of external pressure, is very often responsible for taking the language cycle out of its 'natural' tempo of development (e.g. Milroy 1992b: 204, Thomason 2001). Here, the role of bilingualism and relatively large numbers of bilingual speakers, in particular, has to be emphasised (e.g. Weinreich 1968: 71). Bilingualism serves as a playground for various linguistic phenomena, such as code-switching, borrowing, simplification of grammatical and lexical categories, emergence or increase in periphrasis, to mention but a few. These phenomena may in turn lead to a change (e.g. Thomason 2001: 59-95, Enrique-Arias 2010: 98-9). Numerous studies have confirmed over the years that (extensive) bilingualism can intensify linguistic variation as well as expedite linguistic change at all levels of grammar (Thomason 2001: 10-11, also chapters 5 and 6). Furthermore, the role of acquisition in bilingual communities needs to be considered as well. At this point, a differentiation between bilingual first language acquisition and sequential bilingualism has to be made (Thomason 2001: 49). In bilingual situations, producing deviant structures, in other words, 'imperfect learning' matters, as emphasised by Thomason (2003: 692). She goes on to say that learners are ultimately unsuccessful at acquiring some elements of a source language. These elements are difficult to learn due to their universal markedness (2003: 262 *passim*) (see also below).

With respect to types of contact-induced change, Thomason (2001) offers a broad division into direct and indirect effects. The former involves straightforward importations from a source to a receiving language. The latter can lead to attrition processes or induce a snowball effect, whereby the later changes are triggered by earlier direct adoptions (2001: 61-2). For the second kind of indirect effect, internal pressures may cause subsequent changes and, yet, all the alterations will need to be identified as contact-induced. In other words, unless there is an initial contact-induced change,

the sequence that follows will not occur at all. Furthermore, although anything can be imported by one language from another, the implicational hierarchy of 'word first and grammar later (if at all)' applies to people who spread the innovations as fluent speakers in the receiving language (Thomason 2001: 64). By contrast, for second-language learners of the receiving language the hierarchy is very often reversed (also Thomason 2003: 692).

Apart from potential impact of imperfect learning, there are other factors which create exceptions to the predictability about interference. One important aspect is speakers' attitudes (Thomason 2001: 61, 77, 82). They can either enhance or hinder a change (Thomason 2001: 85). Another factor that very often either induces or hinders contact effects is typological distance (Thomason 2001: 71). Altogether, it has to be remembered that linguistic factors matter less than social ones in a sense that the former can be overridden by the latter (Thomason 2001: 77). In fact, Thomason and Kaufman (1988) regard "the sociolinguistic history of the speakers, and not the structure of their language, (…) [as] the primary determinant of the linguistic outcome of language contact" (1988: 35). Ross (2003) agrees with Thomason and Kaufman (1988) and uses this generalisation in accounting for prehistoric language contact. When using patterns in linguistic data to analyse that kind of contact, scholars are analysing the circumstances in the history of a speech community (Ross 2003: 176). Finally, it has to be borne in mind that although an opportune social setting within the contact situation favours the initiation of a given change, it does not ensure that the change will actually occur (Thomason 2001: 85).

That contact-induced changes differ from intra-systemic changes can be argued in a variety of ways. Still, some of the outcomes in the two types of pressures overlap. Milroy (1984) lists the characteristic changes that occur in a contact environment and subsequently shows how some of these phenomena apply to the changing structures in early English:

> *Certain general principles that operate in language contact situations are now well known to sociolinguists and creolists. These include (1) gross morphological simplification; (2) some loss of segmental phonological distinctions; (3) relexification (i.e. replacement of much of the lexicon of one language – the subordinate one – with the lexicon of the other); and (4) a preference for a fixed SVO word order. ME [Middle English] shows clear signs that at least three of these (1, 3 and 4) had operated: loss of a number of consonantal distinctions seems also to have taken place if the orthography of some thirteenth-century texts is to be trusted (Milroy 1984: 11).*

Items on Milroy's list certainly find parallels in that of Sapir (1921), where he presents drifts which English experienced in the course of its development (previous section). Intriguingly, Thomason (2001) sees the distinction into internal and external pressures as irrelevant, when one concentrates specifically on structural effects of contact-induced change and on the process by which an interference feature is inserted into the system of the receiving language. If there are differences to be distinguished, they ought to be found in the triggers for the change from these pressures. While internally motivated changes are brought about on account of "structural imbalances within a single linguistic system", in contact-induced alterations it is the influence of another

language that constitutes the trigger (Thomason 2001: 86).

And what if externally-motivated interference enhances the already existing problems in structural balance within a linguistic system? Which of the two triggers is stronger? Is there a guarantee a language would drift to the critical point of necessary change without the 'push' from outside? One can attempt to test the validity of these questions when analysing the pace of change of some of the languages in the Germanic family. For example Davis (2006), exploring the Germanic linguistic unity with respect to word-order patterns, asserts that English, together with German and the languages of Scandinavia, "has changed rather more quickly, exhibiting bursts of rapid change" compared with the other languages within the Germanic family, viz. Icelandic (Davis 2006: 155, cf. Thomason and Kaufman 1988: 315). He also remarks that it is difficult to measure the speed of language change and that it is not a constant. McWhorter (2002) places the English language even further off the family matrix. However, contrary to Davis's standpoint, McWhorter's ideas lean on external (Scandinavian impact) rather than internal pressures (McWhorter 2002: 217-72, cf. Thomason and Kaufman 1988: 321).

To continue on the nature of structural effects of contact-induced change, there are several rules and generalisations to be identified. Firstly, it is not likely for the structural interference to occur in isolation, so that any one interference feature is the only one in a receiving language, as observed by Thomason (2001). What is more, in order to make structural interference possible the contact needs to be 'intimate' enough (Thomason 2001: 93). Secondly, structural interference (*metatypy*[45]), along with lexical calquing, as remarked by Ross (2003), are induced by the bilingual speaker's natural impulse to diminish the mental burden by expressing meaning in the same manner in both the primary and secondary lects (Ross 2003: 189). Moreover, in SLA situations, a learner will move to the syntactic communicative mode only in later stages of acquisition. This mode is realised by the sophisticated use of inflectional morphology and of word order as a marker of syntactic relations as well as employment of subordination and exhibiting a fast rate of delivery (Matras 2009: 71). Lastly, with specific reference to changes in word order, these tend to be gradual, much like shifts in non-contact situations. According to Matras (2009), they occur through the "loss of pragmatic specialization of secondary word order variants" (2009: 251). This echoes McFadden's scenario of the loss of applicability of more marked word orders after the loss of inflectional morphology in early English (see section 3.2). In addition, there is the generalization about word order change that places the onset of change at the nominal phrase level. The change affects "more loosely combined elements in the possessive construction" (Matras 2009: 244). It subsequently stirs other attributes, "which behave more generically as attributes, such as lexical adjectives and determiners" (Matras 2009: *passim*). It is only then that the change affects the verb phrase.

The last two points are particularly relevant to convergence scenarios, commonly encountered effects of language contact, whereby languages in contact change in a

[45] Ross (2003) refers to structural interference as *metatypy*. He differentiates his classification of structural interference from that of Thomason and Kaufman (1988), who call it *borrowing*, and from that of Weinreich (1968) [1953], who calls it *grammatical interference*.

manner that makes them alike. Convergence has been identified as a phenomenon that determined the nature of the changes discussed in the present study. Essential features of this phenomenon are discussed in the following subsection.

3.4 MECHANISMS AND MODELS OF PROGRESSION OF CHANGE THROUGH SOCIO-GEOGRAPHIC SPACE

Change in natural languages is universal. It springs from the arbitrariness of language and its conventionality (Keller 1994: 5). Sections 3.3.1 and 3.3.2 listed forces which cause languages to change. Still, to understand and describe the nature of change, a differentiation between the two indispensable components of the process of language change needs to be made. They are innovation (actuation) and propagation (diffusion), as suggested by Croft (2000: 4). The former involves the creation of new items in the language; the latter refers to the manner with which these new items fan out into a broader language usage[46]. To accept innovation and diffusion as two necessary elements in language change means recognising the synchronic and diachronic aspects of linguistic change. Innovation is seen as a synchronic phenomenon, which happens at a particular point in time in speaker action. Diffusion, occurring over longer periods, is a diachronic phenomenon (Croft 2000: 5).

Indeed, innovation begins in the speech of individuals. If they see the innovation as the way to enhance their communicative objectives, the speakers may choose to adopt it into their repertoire and begin to replicate the innovation in new situations (e.g. François 2014: 168). Once the innovation, or generation of variation in the replication process (Blythe and Croft 2012: 271), is carried over (propagated) to other speakers, in other words, selected by other speakers or groups of speakers, the new speech form might settle in, provided it has overcome the competition with the previous norm. It has to be remembered that this is not a question of a single innovation (e.g. Keller 1994) to be tackled but a pool of often competing innovations, and only some of these get selected to be diffused, turning into an actual change. The importance of society's selecting from among the pool of innovations was acknowledged very early on in linguistic studies. For example, Whitney (1979) clearly recognises that it is the community that chooses and changes a language:

> *"Language is not an individual possession, but a social. (…) Acceptance by some community, (…) is absolutely necessary in order to convert any one's utterances into speech. (…) [T]he community [is the] final tribunal which decides whether anything shall be language or not (…)" (Whitney 1979: 149-50, later e.g. Milroy 1999: 23).*

Thus, the propagation of a new variant can be understood as the adoption of a new linguistic convention by the community (Croft 2000: 174). Winter-Froemel (2008), too,

[46] Or the items disappear from a language.

speaks of norms but the perspective is zoomed in specifically on the language system. We can regard change as effected "when linguistic items have acquired a new status within the linguistic system". Otherwise, a complete change is manifested when these items have undergone grammaticalisation or lexicalisation, or when the items have come to constitute a part of the (new) linguistic convention (Winter-Froemel 2008: 240).

Just as the situation with competing innovations is far from simple, so are the patterns of diffusion equally complex. Propagation does not have to target the entire community but may concentrate on a group of dialects, and the subsequent innovations may arise in different parts of the social network (e.g. François 2014: 168). To envisage the multitude of diffusion scenarios, there are different models for spatio-linguistic change. Renfrew (1989) distinguishes four major classes of such theoretical models. Firstly, there is the initial colonisation pattern, which describes language distribution in terms of populating previously uninhabited areas. The second type of model is connected to linguistic replacement, when one language gives way to another (otherwise to other) language(s). The second-last group encompasses theories revolving around the process of divergence, a phenomenon central to historical linguistics, represented in the *family tree*, its most widely discussed model. Finally, Renfrew discusses the category of theories that explain language distribution through convergence (Renfrew 1989: 111). A notable example of a convergence theory is Johannes Schmidt's *wave* model (Schmidt 1872). As the present study deals with the impact of language contact on morphosyntactic changes in early English, this convergence model has been selected as the theory which is expected to help explain the change at issue.

In the wave model, a change begins at one locale, the so-called *focal area*, at a given point in time, and spreads outwards in stages (concentric circles) so that the earlier changes reach the off-lying areas, that is *peripheries*, later (Schmidt 1872: 27[47], Bailey 1973: 69 – point of origin, McMahon 1994: 229, Wolfram and Schilling-Estes 1998: 143, 2003: 713). The wave metaphor is realised in the manner with which particular changes radiate from the central point of contact, diffusing in successive layers, in a way similar to the ripples which spread from the place where a stone is dropped into a pond (Hudson 1980: 41, Aitchison 2003: 43). Although initially designed to describe convergence of languages, a wave pattern can also operate on a more local, dialect level (e.g. Bynon 1977: 193, Bailey 65-109). This model is compatible with scenarios where communities remain in contact. Indeed, it is the linguistic contact that constitutes here a decisive factor in understanding the pattern of language diversification (François 2014: 170). With dialects participating together in the process of innovation, every novel form makes an instance of linguistic convergence (François 2014: 169). The working of the wave is most transparent in descriptions of sound changes, where the examination focuses on an individual item rather than on a more complex pattern. The result of the diffusion of one variant is plotted on a map, along with the path of the spread of other variants creating a network of isoglosses. The theory can also be used to represent the distribution of lexical and grammatical variables.

[47] Currently, scholars rely on an extended and revised model presented by Bailey (1973). Schmidt's account provides a general (concise) description of linguistic differentiation as the progression of waves radiating from the middle point (*mittelpunkt*) (Schmidt 1872: 27).

As with many theoretical models, the wave theory faces some problems when confronted with reality. In practice, a change does not progress so neatly and symmetrically. Diffusion can be constrained by physical (e.g. topographical) as well as socio-political factors (e.g. Wolfram and Schilling Estes 1998: 144, McMahon 1994: 230). As a result, some changes might not spread at all beyond the immediate vicinity of their point of origin. Other developments may be halted almost as soon as they make a start (e.g. Downes 1998: 21). In addition, diffusion may skip a given area on account of various barriers to communication or due to fluctuation in population density. The latter aspect has been, for instance, neatly incorporated into Trudgill's (1974) *gravity* model, also referred to as a *hierarchical* model. His theory accounts for the fact that "changes are most likely to begin in large, heavily populated cities which have historically been cultural centres" (Wolfram and Schilling 1998: 145). From their point of origin, innovations do not diffuse outwards continuously, producing a wave pattern. Instead, the change touches larger urban units first, skipping the nearby, sparsely populated areas. Only later do the innovations spread from more to less populous areas. The logic behind the working of the hierarchical model lies in the fact that larger population units promote greater interpersonal contact, which, in turn, enhances the rate of diffusion of new features. Still, the distance factor has to be included as well. Just as "the amount of interaction between two areas is directly proportional to the population density of these areas, so it varies inversely with the distance between the two locales" (Wolfram and Schilling 1998: 146). Trudgill's model, therefore, factors in the *wave*'s time and distance, adding population density into a pool. However, the theory still fails to account for the role of other social factors such as age, gender and class distinction (e.g. Labov 1966), and for physical factors, e.g. difficult terrain and other natural barriers. What is more, studies by Leslie Milroy (1987) and James Milroy (1992a) have turned attention to even more localised communication networks than the dichotomy between the densely and sparsely populated areas. Their studies revealed that speakers involved with intense and uniform networks are more resistant to adopting new linguistic features than those whose communications expand to other people who belong to different social groups (see Wolfram and Schilling-Estes 1998: 147).

Taking into account all the aforementioned factors influencing the dispersion of a linguistic feature, the wave model appears to be the simplest one. The cascade diffusion, advocated in Trudgill's theory, along with the role of other social and physical factors, has substantial bearing on how changes spread through socio-geographical space. However, the applicability of a given model has to, eventually, be confronted with the quality of the material which provides the basis for the study of the feature. Some of the relevant and differentiating aspects may not be available for factoring (cf. Rissanen 1992: 187-8). Moreover, the relevance of particular factors themselves may also be questioned, especially with respect to the investigated item (cf. Tognini-Bonelli 2001: 66-7).

The present study, uses data sampled from parsed diachronic corpora of early English texts where the time of composition (or copying) of the text as well as its provenance are clearly indicated. Time and distance matter most here, making the wave model the most suitable one for the purposes of this study.

4 Change in Early English Morphosyntax

4.1 OLD ENGLISH MORPHOSYNTAX: STATUS QUO ANTE

Before proceeding with a description of the changes in both early English morphology and syntax which led to a major shift in the identification and realisation of grammatical functions an account of the linguistic situation in Old English and of the conditions preceding the changes, needs to be given.

The standard, and rather simplified, view on the nature of Old English morphosyntax portrays the language as one allowing for substantial syntactic freedom largely due to the availability of inflections. However, upon deeper inspection, the apparent freedom turns out to be rule-governed. While the endings might have, indeed, enlarged the pool of structural possibilities, especially when compared with the syntactic disposition of other languages in the Germanic family, there were quite a few grammatical rules in Old English which had to be observed. An important fact concerning these rules is that they would, ultimately, treat elements regarded as subjects, objects and verbs in a manner visibly different from their present-day representation. For both subjects and objects, the distinction between nominal and pronominal components was of essence, along with the phenomenon of cliticisation[48]. As for the verb element, Old English, much like all Germanic languages, followed a structural requirement whereby the (finite) verb occupied second position in a clause. Lastly, the difference between the main and subordinate clause environments was more pronounced, with specific (distinct) mechanisms used to signal this difference.

To begin with Old English nouns, they had two numbers (singular and plural), four cases in the singular, three in the plural, and different sets of inflectional suffixes for triple-gender, i.e. masculine, feminine and neuter. Next, they would agree in number, gender and case[49] with their corresponding pronouns and modifiers – all of the above constituting building blocks of a noun phrase (NP). Particular cases, in turn, were used to signal the function of an NP in a clause (e.g. Bynon 1977: 148). The nominative (NOM) would be the case of the subject (also of the complement and of address). The case of the direct object was realised by accusative (ACC). It was also used to express the duration of time. The dative case (DAT) expressed the indirect object. It could

[48] The process whereby an autonomous (lexical) item loses its independent properties, turning into a morphological affix, often reduced to the status of a merger with another lexical item (Denison 1993: 476). The development in question frequently has significant implications on word order relations.

[49] It has to be noted that already during the Old English period there was no distinction in form between NOM and ACC. Case syncretism was already under way at that time (e.g. Allen 1995: 24-5). In the plural, nouns were always the same except in the first and second person. In the singular, many of them had the same forms for NOM and ACC and their distinction depended largely on the form of any demonstrative or possessive adjective, or on that of any adjective, which qualified the noun. Next, DAT could be used in the imitation of the Latin ablative absolute. It could also occur after some adjectives, sometimes after comparatives and after some verbs. As for GEN, it could occur after some adjectives and after certain verbs (Mitchell and Robinson 1983: 104-105).

also signal possession as well as some aspects of time. Finally, the genitive (GEN) was used to express the functional relationships between the two nouns which roughly correspond to the present-day English possessive. It could also be used for descriptive purposes and to show partitive relations. It was also not uncommon for GEN to realise some adverbial functions (Mitchell and Robinson 1983:104-105, Denison 1993: 21, Fischer and van der Wurff 2006: 110). Depending on the function, both NPs (whether subjects or objects) as well as verbs tended to occupy different (particular) positions in the clause (Fischer and van der Wurff 2006: 110 and 156).

Nominal function on the surface, as noted above, was predominantly indicated by case. There were, nevertheless, other means of designating syntactic roles. One was the use of prepositions; the other, expectedly, by means of word order. With respect to the former, certain underlying structures were typically associated with prepositions. In most scenarios, prepositions were optionally selected as an alternative to the noun-case preference. There would be situations, however, where the structures could not do without prepositions, such as passive environments (i.e. an overtly expressed agent in a passive sentence had to be realised by a prepositional phrase). In addition, prepositions would impose particular cases on the nouns with which they formed a construction. On one hand the selected case was the same as the one chosen by the underlying function; on the other, they would not correspond to each other. Last but not least, certain verbs could override the rules of case assignment and impose their own case selections (e.g. Traugott 1972: 80-81).

As regards OE word order, taking subjects and verbs as the most vital building blocks of a sentence, the following ordering combinations can be distinguished in the prose, following Mitchell and Robinson (1983: 61):

(i) SV with the verb immediately following the subject,
(ii) S ... V with other elements of the sentence coming between the subject and the verb,
(iii) VS, where the subject follows the verb.

As in Modern English, the word order in (i) occurred in main as well as in subordinate clauses. In this respect, the two clausal environments may have been indistinguishable. Among variations of this order (also found in Modern English) were cases when an adverb preceded the (main) verb as in: ic hæbbe nu gesæd hiora ingewinn (I have now told their intestine-struggles) (Traugott 1972: 94). Furthermore, as in Modern English, the indirect object would precede the direct one. Other syntactic peculiarities were more characteristic of the language of earlier days. Notably, the negative particle ne was placed before the verb. This rule came to be observed in all three word orders. Pronominal objects were preverbal while nominal objects tended to follow the verb. Lastly, in cases of complex verb phrases, an infinitive or a participle appeared in clause final position (e.g. Traugott 1972: 107).

The order in (ii) was characteristic of subordinate clauses (e.g. Koopman 1998: 141). However, it was not uncommon to see this syntactic arrangement also in the main clauses (e.g. Koopman 1998: 142, Stockwell and Minkova 1991: 381). The position of the verb was not exclusively final either. Further, the cases where the adverbial ex-

tension would follow the verb were frequent: *ær he acenned wæs of Marian* (before He was born of Mary) (Mitchell and Robinson 1983: 61). The 'subordinate order' started to decline during the late OE period and by later ME it was no longer present.

The third ordering combination, which in present-day English still operates in questions with the verbs BE and HAVE, such as *Are you happy?*, was typically found in OE positive non-dependent questions, regardless of the presence of interrogative words. It would also appear in negative non-dependent questions, such as: *ne seowe þu god sæd on þinum æcere?* (did you not sow good seed in your field?) (Mitchell and Robinson 1983: 62). In addition, VS was found in positive statements as well as in some negative statements. Finally, the subject would follow the verb in subordinate clauses of concession and condition; and in main clauses introduced by certain adverbs.

Word order was, indeed, an important marker distinguishing matrix from subordinate clauses (e.g. Fischer *et al.* 2000: 88-9). However, as previously indicated, order variations were possible and, at some point, other clues needed to be available to make the distinction between clauses clearer (cf. Koopman 1998: 141). An effective signal indicating the main clause environment was the presence of topicalisation and verb-fronting – both consonant with the verb-second constraint (Stockwell and Minkova 1991: 385 and 379, Koopman 1998: 135, cf. Fischer and van der Wurff 2006: 182). Signs of subordination were doubling of the correlative (e.g. *þa þa*, *swa swa* etc.) and use of the subjunctive (Fischer *et al.* 2000: 89, Stockwell and Minkova 1991: 373). Finally, a large part of the nature of the clause could be inferred from the context (Koopman 1998: 141).

A closer look at Old English word order patterns reveals the importance of the distinction between the finite and nonfinite elements in the verbal group. The position of the former at that stage of development of the language was especially paramount for the matrix clause level, where it would very often come after the first constituent (regardless of its function), invoking the so-called verb-second (V2)[50] constraint, as in: *þā ferde se biscop...* ('then went the bishop...') (Stockwell 1981: 576). In addition, it is important to note the type and position of the subject within the context of topicalisation. In this kind of context, pronominal subjects tended to appear before the finite verb whereas full NP subjects would follow it (e.g. Fischer *et al.* 2000). The phenomenon of V2, characteristic of all languages in the Germanic family, ceased to be a prevalent structural requirement in English a few centuries later, towards the end of the Middle English period (around 1500) (e.g. Swan 1991: 235). Interestingly, the rest of the Germanic family has kept the V2 constraint active until today (Swan 1991: 233,

[50] A distinction between CP-V2 and IP-V2 languages needs to be made and the relevance of this difference will be of essence especially in section 4.3. All V2 languages show V2 in main clauses. Still, the CP-V2 language subtypes exhibit the verb-second order in the embedded clauses which in some respect have the structure of matrix clauses, whereas the IP-V2 subtypes show V2 in a broad range of subordinate clauses (Kroch *et al.* 2000: 355).

Fischer *et al.* 2000: 82)[51]. Since the increased absence of V2 (in ordinary declarative clauses) in early English is regarded as one of the primary reasons for the emergence of the rigid SVO order – the focus of the current study – it will be treated in greater detail in section 4.3.1.1.

There is so far no clear consensus about the word order status of Old English. Some scholars see Old English as an SOV language drifting towards SVO (and verb-second), consider e.g. Robinson (1992: 166). Then there are others like, for example Denison (1993), who are undecided whether we should regard OE as an SVO or SOV language (1993: 28). There are also proponents of the view that Old English was clearly SOV (e.g. Traugott 1992: 274). However, studies exist showing that SVO, too, was fairly well established in the latter part of the period (Swan 1994: 235). What is more, the verb finality in subordinate order, similar to the already mentioned V2 constraint, was never categorical in Old English (e.g. Stockwell and Minkova 1991: 374). Much of the debate on the word order of Old English is, indeed, connected with different views on the character of particular constituents in the earliest English morphosyntax. However, it also impinges greatly on the subsequent syntactic changes, where the gradual dominance of SVO emerged, undoubtedly, as a result of a combination of different factors, among them the loss of inflections. This was followed by growing prevalence of SV, a shift from OV to VO, and the loss of the validity of the verb-second constraint. The following sections treat in greater detail those factors that led to the emergence of the subject-verb-object order.

4.2 MORPHOLOGICAL SIMPLIFICATION IN EARLY ENGLISH

Although the underlying functions of NPs remained the same for both Old and Middle English, their surface representation changed fundamentally during the latter, ME period (e.g. Traugott 1972: 121, also White 2002: 124). All case marking, except for GEN, was lost for nouns and significantly reduced for pronouns. The collapse of the ACC vs. DAT distinction in nouns was the first major step. Next, DAT was first generalised, only to be completely lost soon after. The loss of the distinction left nouns unmarked on the surface, apart from GEN (Traugott 1972: 122). As for pronouns (and determiners), the DAT and instrumental (INSTR) cases merged into one, with the shift concluded by 1200. In addition, the DAT-ACC contrast was lost and the DAT option generalised for most pronouns. In this respect, there was eventually a division into only three: nominative, genitive and oblique (object) cases. The ACC was generalised in one setting – the third person singular neuter form *(h)it*. The form for NOM and ACC was the same. It made *(h)it* act like a noun with only genitive and unmarked forms. Traugott refers to the process as "the operation of functional load" (Traugott 1972: 122).

---

[51] V2 has applied to the rest of the Germanic family virtually without exception (Fischer *et al.* 2000: 82). By contrast, Old English V2 would be optional (Swan 1994: 236). According to Stockwell (1981), V2 of OE was a word-order norm that was not fully grammaticalised (1981: 576). Moreover, although no longer a full-fledged V2 language (V2 in ordinary declarative clauses), Modern English does retain vestiges of V2 constraint in certain contexts, e.g. sentences with a non-subject negative phrase, which require a modal or a form of *do* in the second position (Santorini and Kroch 2007).

As the process of 'natural gender' taking over the grammatical gender unfolded, there arose a need to clearly mark the contrast between male and neutral in pronouns. Indeed, Curzan (2003) stresses the role of (personal) pronouns in the change of the gender system in English (2003:19). However, there were other contributing factors, as noted by Bertacca (2009). The levelling of the inflectional endings on nouns must have played an important role too. The simplification of the inflections of the Old English demonstrative pronouns (with the consequent development of the definite article) as well as the erosion of adjectival endings were also significant in the establishment of natural gender (Bertacca 2009: 55). As the early morphosyntactic conditions were changing, the functions no longer represented by cases were taken over by word order and prepositions. With respect to the latter, prepositional constructions, already available in OE, began to be utilised more extensively in ME in order to compensate for the lost functional aspects previously realised by cases. Interestingly, of all the cases, only GEN has remained fully functional, although not without some restrictions (Traugott 1972: 122-3). Finally, the inflectional loss on the verb phrase was nearly as drastic as on the noun phrase. The OE distinctions – three persons, two numbers (sing. and pl.), three moods (indicative, subjunctive and imperative) along with two infinitives (*casus rectus* and *casus obliquus*) were all discarded. The only inflection left was (and still is) the *–s* ending for the marking of the third person singular (indicative) (White 2002: 124).

When discussing the beginnings of morphological erosion in early English, a condition needs to be set with respect to the scale as well as the material which provides evidence for the change in question. In the present study, the starting point is the moment when the petering out of inflections was detectable in the early English texts.

It is a well known fact that morphological simplification began to be most visible during the early Middle English period first in the North (e.g. Milroy 1992b; Fischer 1992). From there the tendency proceeded gradually southwards through the Midlands and finally reaching the South (e.g. Millward 1996: 142, cf. Fisiak 2005: 60f). Texts such as the Peterborough Chronicle, which greatly aligns with the morphological conditions of the North, show that the loss of endings was much more advanced there, with the concomitant strategies of making up for losses by means of prepositions and word order following suit (e.g. Traugott 1972: 122-3, Görlach 1997: 62).

As regards the timeline for the loss of inflectional variation within NPs, most of the reductions along with the spread and establishment of 'new' morphological tendencies encompassed the period of approximately six hundred years (Figure 4.2a). To begin with the NOM-ACC distinction, the syncretism of the two cases was well-advanced by the end of the tenth century. It would most certainly be completed for the North and the northeast Midlands before the end of the Old English period (i.e. the twelfth century)[52]. The rest of the dialects underwent the change during the Middle English era, with the developments finally reaching SW and Kent about 1400 (cf. Hogg 1997: 101, Fisiak 2004: 73). Secondly, the 'zero' DAT sing., before spreading southwards, began to operate in the North and parts of the Midlands also by the end of twelfth cen-

[52] The NOM-ACC syncretism advanced more rapidly in the plural forms than in the singular ones, even in the most conservative dialects of OE.

tury. Similar to the NOM-ACC pair, it became dominant in the South by the beginning of the fifteenth century. Thirdly, the strong, masculine *–es* GEN ending, replaced, and in some environments disambiguated, all other older forms in the course of the twelfth to fourteenth centuries (Görlach 1997: 62)[53]. Further, the extension of the *-es* plural formative, predictably, proceeded on a larger scale first in the North (1150-1200) and the East Midlands (1175-1225) (Newman 2008: 116). In the South, the ending began to prevail a little later due to the competition with another plural marker *–en*. The *–es* plural finally started to dominate throughout as late as the second half of the fourteenth century (Fisiak 2004: 74).

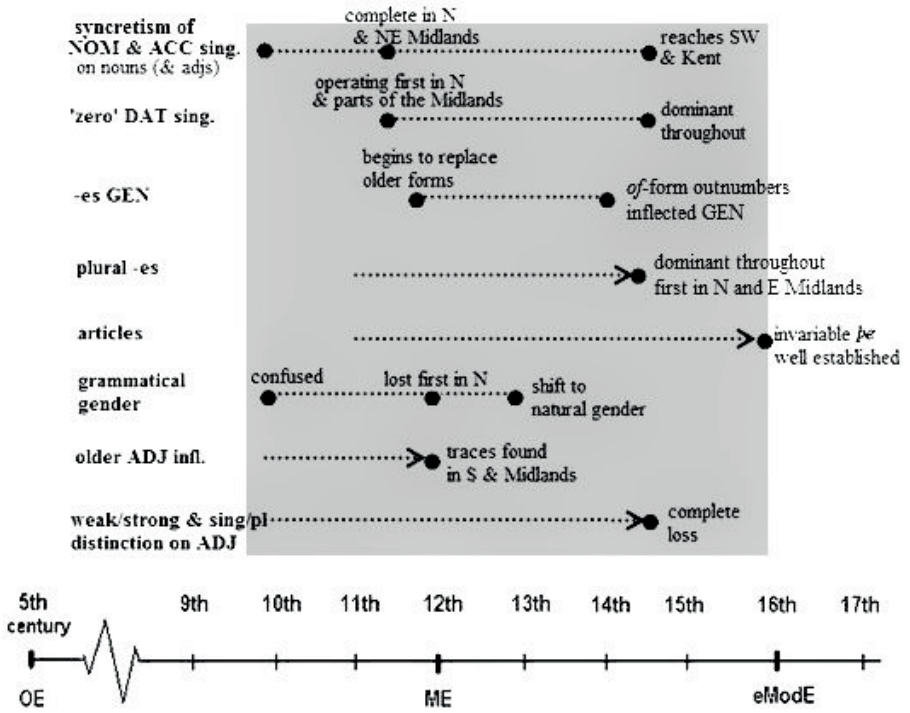

Figure 4.2a Timeline of the loss of inflectional variation within NPs

Grammatical gender, which had its share in providing a wide array of inflectional suffixes, was lost quite early, first in the North. The tenth century texts show that the distinctive endings *–es*, *–re*, *–ne* (GEN, DAT, ACC sing.) applied to all attributive articles and adjectives regardless of gender. More often than not, the concept of gender must have appeared rather vague, which made it invalid in some dialects as early as by the twelfth century. The shift to natural gender occurred a century later (Görlach

[53] According to Allen (2002), the reanalysis of GEN on nouns as a clitic occurred during late 14th century (2002: 75). As for the competition between the incoming *of-* form and the inflected GEN, the outline provided by Szmrecsanyi (2013) shows that while the inflected GEN greatly outnumbered the *of-* form until the 12th century, by the 14th century the situation was reversed. Interestingly, there is a revival of the popularity of the *s*-genitive during the Early Modern English period (Szmrecsanyi 2013: 61).

1997: 61-2, Lass 1992: 107). With respect to adjectives, traces of the older inflections still lingered in the South and the Midlands at the beginning of the Middle English period. The complete loss of adjectival endings took place before 1400, when the final /–ə/ was dropped, removing the weak-strong and singular-plural distinctions (Fisiak 2004: 78). Finally, the number of forms of the article was significantly reduced, as expected, first in the North (White 2002: 158). The invariable *þe* became established by 1500 (Görlach 1997: 62).

As for (personal) pronouns, the well-developed OE system experienced radical simplification during the Middle English times, welcoming the emergence of the single, uniform 'object' (i.e. oblique) case. Many of the novel pronominal forms had their own pace and place of occurrence (e.g. Fisiak 2004: 80-3, Görlach 1997: 64-7). Nevertheless, it is worth mentioning that the falling together of INSTR and DAT occurred earlier (completed by 1200) than the loss of the ACC-DAT distinction. In the case of the latter, it is the DAT form that was generalised for most pronouns (e.g. Traugott 1972: 122-3). Next, the dual forms were replaced by the plural sometime around the thirteenth century (Görlach 1997: 65). What is more, some of the ineffectively differentiated aspects of third person pronouns were refined during the course of ME. Many of these aspects, notably the feminine singular *she*[54] and the distinctive plural form *they*, appeared first, again, either in the North or in the East Midland area (Millward 1996: 146, Krygier 1998: 117-124, Manfred 1995: 166, cf. Britton 1991[55]). Lastly, it has to be remembered that the early system of personal pronouns was linked with that of personal endings on verbs.

In verbs, the picture of changing conditions is even more complex, as the alterations concerned both the inflectional and thematic aspect of the verbal environment. For the simplification in the inflectional system, the pattern of reduction is just as varied as in the case of pronouns. However, some peculiar regional preferences can be distinguished at the onset of the levelling of numerous forms. The dropping of the *–e* ending (singular *present/preterite*[56] indicative/subjunctive) occurred first in the North at the end of the twelfth century, reaching the South only two hundred years later. Similarly, in the North the 'zero' ending was generalised in all persons (excluding the second) present indicative when directly followed or preceded by a personal pronoun, as in: *say wē* (Fisiak 2004: 86; Fernandez-Cuesta and Rodriguez-Ledesma 2006: 102). Next, the *–en* ending, designating the infinitive, plural *present/preterite* indicative/subjunctive and the past participle, was abandoned first in the North as early as the beginning of the Middle English period. The tendency arrived in the South at the end of the fourteenth century. In addition, the second person preterite indicative *–est* was discarded earliest in the North as well. The development percolated through the Midlands and subsequently the South much later, before 1400 (Fisiak 2004: 87).

---

[54] The characteristic northern form: *sco* is attested in the early 14th century (Fernandez-Cuesta and Rodiguez-Ledesma 2006: 105).

[55] Unlike some scholars who speak of ON (Curzan 2003: 192) or Celtic (Ahlqvist 2010: 63-8) influence, Britton (1991) regards *she* as a native development. His view is popular though not unanimously accepted (Curzan 2003: 191).

[56] The distribution of *preterite* forms on a dialectal map was different from the forms of the *present*. There was a clear line of demarcation between the North and the rest of the dialects (e.g. Fisiak 2004: 87).

Two facts can be established from the outline drawn above. The first is that the most intense phase of morphological simplification is reserved for the period between the end of OE and the middle of ME. The second is that the geographical patterns for the majority of morphological simplification scenarios observed in early English are virtually identical (Figure 4.2b).

Figure 4.2b Path of morphological change on a dialectal map of early English

The northernmost dialects were the most innovative (cf. Trudgill 2010: 33). Numerous levelling tendencies were initiated up in the North. From there the alterations spread with variable speed to the Midlands before finally penetrating the South (e.g. Fisher 1992: 208, Millward 1996: 142, Tristram 2002: 125). The areas of the southeast remained the most conservative. They would very often resist the changes the longest, preserving the features which had been discarded centuries before in the North (e.g. White 2002: 157-8).

4.3 SYNTACTIC SHIFT IN EARLY ENGLISH

4.3.1 Rise of SVO in Middle English

The emergence of a strict SVO order, a syntactic framework where grammatical relations are expressed by specific, fixed position of constituents in a clause (Fischer and van der Wurff 2006: 114) rather than by means of overt morphological marking, was a complex development, incorporating various small-scale rearrangements. The preference for this particular pattern began to transpire during the Middle English period,

taking several subsequent centuries to bloom and prevail. Its advent might be viewed in terms of occurrence of secondary modifications around particular constituents of the clause, notably the growing pervasiveness of the overt SV configuration and the shift from an OV to a VO pattern. Other developments occurring alongside the changes mentioned above included the changing characteristics of pronouns. In Old English they appear in different positions from those occupied by full NPs. During the Middle English period the positional asymmetry (van Kemenade 1987: 138-40) between the two is removed, which has serious consequences for the realisation of both subjects and objects as clausal constituents. All subjects whether pronominal or nominal now come before the verb and usually occupy the first position in a clause. Pronominal objects, too, start to follow the distribution of full NPs in that function. Last but not least, a phenomenon linked to the changing status of pronouns and subsequent word order shift was the loss of the validity of the V2 constraint. This syntactic rule prevailed in Old English and continued to operate well into in Middle English. However, in the later ME period a dialectal discrepancy between the North and the South with respect to realisation of the V2 constraint developed, which eventually impinged on the applicability of the rule itself. In texts from this period, the number of sentences with a 'transparent' V2 constraint began to gradually decrease from the end of Middle English onwards (Fischer *et al.* 2000: 184-5).

As pointed out above, the Middle English period witnessed the gradual establishment of the general characteristics of the order operating in present-day English – Subject (Auxiliary) Verb Object (Traugott 1972: 160, Hock 1986: 332) or verb-medial (Bech 2001: 70). The alterations which occurred at the turn of the Old and Middle English eras, in particular the loss of case distinctions (e.g. Robinson 1992: 166; Allen 1995: 417) as well as the erosion of verbal endings, are often mentioned as a significant factor in establishing a rigid SVO order. The lack of overt distinguishing morphological markers is thought to have created problems in the identification of clausal constituents, notably NP subjects and NP objects (e.g. Hock 1986: 369). The distinction between objects themselves is said to have been disturbed too (Fischer and van der Wurff 2006: 189), which then led to the loss of semantic contrast within the verb-object relations (e.g. Fischer and van der Wurff 2006: 164-5). Moreover, it will be shown that the erosion of inflections was instrumental in the changing properties of personal pronouns. Finally, the loss of verbal inflection in particular is thought to have contributed to the ultimate demise of V2 in early English (e.g. Haeberli 2002: 89 on the loss of empty expletives; Kroch *et al.* 2000: 353-91 on competition of the original V2 grammar with a non-V2 grammar in contact situations). It would appear that the loss of inflectional morphology in the transition from Old to Middle English is the underlying cause for generating difficulties around the recognition and processing of the functions of the main syntactic units and for creating problems in the applicability of major syntactic rules. In this respect, the subsequent growing prevalence of a strict word order could be regarded as a countermeasure against potential ambiguities and inconsistencies around the semantic roles of both pronominal and non-pronominal noun phrases. SVO ordering also stepped in when non-categorical constraints, such as the V2 rule, became increasingly incompatible with the arising syntactic conditions. The following sections discuss explanations for the growing prevalence of SVO in early English.

4.3.1.1 Emergence of overt SV, pronouns, V2/V3

The increased overtness of the subject and its firm placement before the verb (SV) was one of the major development on the way to strict SVO. Firstly, back in Old English the subject could, but did not need to be clearly expressed in the clause, and the presence of inflections facilitated variability in that respect (e.g. Traugott 1972: 84 and 131). The phenomenon is especially frequent with pronoun subjects[57], as in *Gemunu þā mǣla* ('[I] remember the occasions'). It was natural in successive sentences, as long as there was no ambiguity (Mitchell and Robinson 1983: 106; 2011: 101, 257 n.212). By the end of the ME period that option, however, was lost from the language altogether (Fischer and van der Wurff 2006: 168). With the loss of surface grammatical markers, the bare NPs appearing preverbally came to be interpreted as subjects, as suggested by Hock (1986: 369) and also by Fischer and van der Wurff (2006: 173). This was, in fact, an extension of an already existing option, as the leftmost NP of a sentence was usually a subject (Traugott 1972: 130, Harris 1981: 185).

Furthermore, a shift in the functional status of personal pronouns additionally enhanced the SV tendency and questioned the applicability of the V2 constraint in early English. In Old English, pronominal subjects and objects were different from full NPs. They were clitics rather than independent constituents and appeared in preverbal position. Stockwell and Minkova (2000) suggest that this particular status made pronouns syntactically invisible, assigning them morphological properties. Thus, their status is identified as the one '"outside" the normal constituent-order rules' (2000: 289) Only by treating personal pronouns as clitics can we speak of the V2 constraint, as observed for all the other Germanic languages (also in Holmberg and Platzack 1995: 63). The following examples would otherwise be violations of the constraint:

(2) *he* him hæfde geseald *gislas ond aðas*
('he to-him had given hostages and oaths')
(3) *7 him þa gislas sealdon*
('and to-him the hostages they-gave')[58]

The special properties of personal pronouns can be seen expressly in the behaviour of object pronouns, which followed their subject counterparts. As for the latter, the frequency of preverbal pronouns subjects made scholars identify the so-called "verb third" (V3) order in main clauses, as in:

(4) Ælc yfel he mæg don
Each evil he can do[59]

[57] Unexpressed objects were a possibility as well, as in: *ac ic come and fullode on wætere,* ('but I came and baptized [him] in water,') (Mitchell and Robinson 1983: 81).
[58] Examples taken from Stockwell and Minkova (2000: 289).
[59] Example taken from Kroch (2008: 714).

Thus, there would be an alternation between V2/V3, regardless of whether pronouns would actually exhibit the clitic behaviour.[60]

During the ME period, in northern dialects, subject pronouns began to lose their clitic status and were subsequently treated like any other subjects (Fischer and van der Wurff 2006: 184). This, in turn, allowed invoking the V2 rule quite consistently, which as observed above, would normally not operate with pronominal subjects. Lightfoot (1999) proposes that the process of decliticisation of pronouns in the north was a result of heavy influence from Scandinavian - a language with persistent V2 and non-clitic pronouns. Speakers of varieties influenced by Scandinavian language(s) would be prone to interpret the southern constructions with clitic subject pronouns as SV... clauses without inversion. Such a course of events, as triggered by language contact, was further propagated by increased mobility of the population (Fischer and van der Wurff 2006: 185).

Stockwell and Minkova (2000) also discuss the possibility of Scandinavian influence in the process of decliticisation of personal pronouns but they show more restraint about the volume of the foreign interference. Apart from potential impact from ON, they explore two other hypotheses. One of them is connected to the lack of consistency in the primary linguistic data which may have led to incorrect abductions about archaic pronominal positions made by a child. The second hypothesis puts the drift to SVO to the foreground, as the analogical basis for the shift in the position of pronouns (2000: 297). They present several factors which could be instrumental in the loss of the clitic status on pronouns. Unlike Lightfoot's scenario presented above and the earlier proposal of van Kemenade (1987), which focuses on subject pronouns, the loss of verbal inflection and the V2 constraint, Stockwell and Minkova's ideas concern object pronouns and the factors they raise are related more to the breakdown of the case system. The first is connected to the emergence of the ACI[61] construction with 'expect' type-verbs in early Middle English, which in OE was considered foreign (Stockwell and Minkova 2000: 299) The second factor concerns the syncretism of pronominal forms, especially the loss of contrast between 3rd person ACC and DAT pronouns. The third factor involves the survival of the so-called "quirky" subjects[62] (2000: 300). In all these cases, the morphological changes made the acquisition of pronominal clitics more complex, if not too difficult. Learners would be faced either with the abundance of new or the absence of previous options around object marking. Finally, Stockwell and Minkova (2000) identify the analogical pressure from the generalisation of SVO order as an impulse activating the decliticisation process. The authors entertain the idea that the shift to SVO may in fact be the main cause for postverbal placement of pronouns and that no other theory needs to be formulated (2000: 301).

---

[60] Koopman (1992) identifies the non-clitic properties of some of the Old English pronouns (1992: 61-3). More recently, Bech (2001), too, emphasises the fact that some of the subject pronouns cannot be regarded as clitics, even though they occupy the same position as the clitic elements (2001: 98). More details for the 'un-Germanic' V2/V3 alternation in OE are found in paragraph below.

[61] I.e. Latin-type accusative and infinitive constructions (Fischer 1994: 91).

[62] I.e. non-NOM NPs with syntactic properties of subjects (e.g. Allen 1995: 3).

The manifestation of the so-called V3 order mentioned earlier, where a verb follows both a fronted constituent and a subject pronoun, presents some problems with respect to generating parallel analyses of the typically Germanic V2 syntax. Old English V2 is different from that of modern Dutch or German. Clearly, subject-verb inversion will apply uniformly in all of these languages in operator-fronting contexts such as in questions, in negative clauses, and in clauses containing some adverbs which are not operators[63]. However, the OE V3 case presented in example (4) above would be unacceptable in other Germanic V2 languages, which would still employ the V2 constraint in this case (e.g. Haeberli 2001: 202). What is more, while for non-pronominal subjects, the fronting of a non-operator in a main clause always generates V2 in modern Dutch in German, it would not necessarily do so in Old English. Thus, there would be examples such as:

(5) *On his dagum sende Gregorius us fulluht*
In his days sent Greogory us baptism
'In this time, Gregory sent us Christianity'

Here, V2 is, indeed, observed. However, one could also find instances like the one in (6) below:

(6) *ðone Denisca leoda lufoað swyðost*
that Danish people love most
'The Danish people love that one most'[64]

There is, again, a possibility of V3, which would be ungrammatical in both Dutch and German. Interestingly, the non-categorical use of V2 notwithstanding, this constraint is still much more frequent than V3 in OE in such contexts (Haeberli 2002: 90).

While the Early Middle English syntactic tendencies resemble those of OE, the situation changes during the later period and by 1400 there is a visible receding of instances with the V2 constraint (Haeberli 2002: 91). Various proposals offered to explain the loss of validity of V2 in early English can be divided into two groups, as suggested by Haeberli (2002). One of them represents scholars who advocate the loss of V2 as a result of decliticisation of subject pronouns (cf. Fuss 1998, van Kemenade 1987, Platzack 1995: 200-26, Roberts 1993[65]). Their ideas boil down to the assumption that the loss of the clitic status of pronoun subjects leads to the "assimilation of the two types of subjects" (Haeberli 2002: 92) which, in turn, erases the V2/V3 contrast, leaving the V3 order as the only operational option. Haeberli (2002) raises three objections to the decliticisation hypothesis. One of them reveals that the decliticisation apparently creates the opposite effect to the process described above. Following Lightfoot (1995: 49, note 3), Haeberli points out that with the loss of asymmetry between pronominal and

[63] These include instances such as *þa* (then), *þonne* (then) and *nu* (now), as in: *þa ge-mette he sceaðan* ('then met he robbers', ModE: 'Then he met robbers') (Haeberli 2001: 202).
[64] Both examples taken from Haeberli 2001: 203).
[65] All the references found in Haeberli (2002: 89).

non-pronominal subjects the former would rather be expected to actually begin to appear in V2 contexts, much like the latter. This possibility was already mentioned in the earlier paragraph where the North-South dialectal discrepancy in ME with respect to the presence of clitics was presented (Lightfoot 1999). The second problem concerns the fact that the explanations for the loss of V2 through decliticisation focus directly on the end result and do not really attempt to answer the question why pronouns would lose the clitic status in the first place. Finally, the decliticisation scenario could prove problematic, according to Haeberli, due to the availability of evidence against the stage in ME whereby pronouns began to behave like full NPs. There are studies which show that object pronouns exhibit certain properties which make them markedly different from full NP objects after the beginning of the 15th century (Allen 1995: 419, Haeberli 1999: 400, 412 ff.). This leads Haeberli to conclude that if decliticisation occurred in ME, object pronouns were not involved. However, studies such as those of Stockwell and Minkova (2000) outlined above show that the strength of the interaction of the syntactic and morphological components cannot be overlooked. This is also emphasised by Anderson (1993: 38), and decliticisation of objects has to be considered too (also McFadden 2005: 63-82).

The other group of proposals explaining the loss of V2 in early English is represented by proponents of contact between different grammars as the impulse behind the change (e.g. Lightfoot 1997; Kroch *et al.* 2000). These authors speak of potential dialect variation between the South with (IP)V2 where the pronominal/non-pronominal subject contrast was distinguished and the North which had the (CP)V2 pattern applying uniformly to both types of subjects. They argue that the loss of validity of V2 is a result of contact between two dialects. Haeberli (2002) stresses the fact that although plausible, the grammar contact explanation lacks clear empirical evidence for or against it (2002: 93). Here, he puts forward his explanation for the loss of V2 in early English. According to Haeberli, the V2 constraint withers due to the loss of the licensing of empty expletives, the process correlated with properties of the verbal agreement morphology. The loss of licensing was caused by the erosion of the infinitival *n*-ending in late Middle English. It became identical to the first person singular ending at which point the licensing condition is not met, as it can occur only when these forms are clearly differentiated (2002: 102-106). The idea presented by Haeberli, thus, connects the loss of V2 with changes in the inflectional morphology.

Much like with the emergence of the overt SV, the demise of V2 and the changing behaviour of personal pronouns, the loss of inflections will be implicated (among others) in the growing preference of order where an object invariably follows the verb, which is the theme of the following section.

4.3.1.2 From OV to VO

Another syntactic development contributing to the reinterpretation of rules, prompting the move towards rigid SVO in early English was the shift in object placement, from a (typically OE) preverbal (Fischer *et al.* 2000: 99) to a (ME and onwards) postverbal position (Fischer *et al.* 2000: 162). That said, VO was not at all uncommon in Old English. It would appear especially with nominal objects (Fischer *et al.* 2000: 161). At the turn of the linguistic epochs, during the eleventh and twelfth centuries,

there was a sharp increase of surface VO order (Fischer *et al.* 2000: 178). Still, postverbal objects began to outnumber the preverbal ones only around 1300, i.e. half-way through ME and became the norm by the end of the period (from 1450) (Fischer *et al.* 2000: 162). What is more, the Middle English OV would still allow not only pronouns but full NPs as well.

With respect to the earliest instances of VO order, two approaches need to be considered here, as pointed out by Fischer and van der Wurff (2006). The first approach is advocated by van Kemenade (1987), who argues that VO arrived as a result of the operation of extraposition, during which the object was moved from preverbal to clause-final position. The second option is offered by Pintzuk (1998). She claims that since OE allowed for VO as well as for OV, competition between these variants could be a likely scenario. Though insightful, these alternatives do eventually seem to be lacking, as remarked by Fischer and van der Wurff. Extraposition might have worked at the beginning. However, in due course the VO sequence becomes a full-fledged word order, no longer driven (if at all) by secondary developments (cf. Stockwell and Minkova 1991: 377). In my view, the second approach points to the right direction. Still, according to Fischer and van der Wurff, the development was very gradual and comparing only single texts might not provide us with a complete picture (2006: 186).

As for the increase in surface VO order in early English, there are several early yet insightful proposals. Firstly, the emergence and stabilisation of VO is attributed to the existence of afterthought-like elements that appear postverbally, as proposed by Stockwell (1977: 291-314). These include relative clauses, conjuncts (split constituents), appositions, and adverbs. Faced with preponderance of these structures, language learners would infer that VO was a general norm.

The second proposal rests on the idea that the postverbal position was generalised from heavy to lighter objects. The position would at first have some expressive value which would become bleached because of common use (e.g. Lightfoot 1979: 393, 1981: 231) or due to dialectal diffusion (van der Wurff 1990: 42-4).

Thirdly, the change from OV to VO might have been due to particular perception problems caused by the use of OV order for internally complex objects (Colman 1988: 33-89). More recently, argument for the shift from OV to VO, related to "interaction between the evolution of relative clauses and perceptual factors", is pursued by Ogura (2001: 243).

The fourth scenario promoting the emergence of VO is based on the infrequence of OV in the main clauses (Lightfoot 1991). According to Lightfoot, these clauses are the prime environment guiding learners (Lightfoot 1991: 56-76). His proposal is further commented on in the next section (4.3.2).

Fifthly, VO order is claimed to have been reinforced by the loss of overt case marking on noun phrases (NPs), as advocated for example by Weerman (1989: 157-78) and Kiparsky (1997: 460-94). For Weerman, the loss of case inflections eradicates the possibility of alternation between OV and VO in Germanic languages including English. Once the case inflections disappear, individual languages have to settle on one of the two options. For Kiparsky, structural position and inflectional morphology are alternative arguments licensers. From these he derives the generalisation that loss of inflectional morphology involves fixed order of direct nominal arguments.

Roberts (1997) provides a sixth scenario for the causal relationship between the rich inflectional morphology and the change from OV to VO. Unlike some of the earlier theories provided by for example van Kemenade (1987) or Pintzuk (1991) where the word order shift is analysed through a change in the directionality of VP, Roberts's account realises changes in word order through the loss of object movement (1997: 397-426). Though insightful, his ideas have been subject to criticism, notably by McFadden (2005)[66]. The main objection revolves around the fact that the theory proposed by Roberts does not agree with the facts of the history of English and that his ideas also go against the evidence from the other Germanic languages (2005: 63-82).

Finally, the emergence of VO at the expense of OV has also been attributed to language contact situations. Of the many possibilities potential impact from Scandinavian has often been mentioned as being instrumental in propagating the use of VO. The impact of ON was already mentioned when discussing the demise of the V2 constraint (section 4.3.1.1 and the discussion on the contribution of Kroch *et al.* 2000). Proposals suggesting Scandinavian influence on early English structures will be discussed in detail in section 4.4.2 However, it is worth mentioning at this point the work of Weerman (1993). Much like Kroch *et al.*[67], he comments on the role of acquisition in the process of linguistic change. For Weerman the interplay of first and second language acquisition is the key to understanding some of the changes occurring in a language (1993: 903). Thus, according to Weerman, the V2 of OE allowed for many contexts where the underlying OV order was overwritten by surface VO. The tendency was enhanced by the linguistic behaviour of the resident Scandinavian population. L2 learners, i.e. Scandinavians with English as their L2, would overgeneralise VO to cases where OV would be more fitting. If the L2 performance constituted the primary linguistic data in which children were immersed during the learning process, the next generation would have already had VO as their default order. As for the role of contact in the decline of OV, Fischer and van der Wurff (2006) mention the possibility of direct influence from French. Nevertheless, according to them the contact account can hardly explain the loss of preverbal pronominal objects, as the southern texts exhibit more OV than their northern counterparts (2006: 187). One would expect the reverse when tracking French influence in the written data.

4.3.2 Morphosyntactic change and clause types

Much of the debates on the emergence of SVO in early English revolve around the question of which of the clause environments, main or subordinate clauses, is the most likely source of the change in question. The traditional view holds that the word order developments in main and subordinate clauses occurred independently (cf. Stockwell and Minkova 1991: 400). However, many recent studies indicate that the two environments could be connected, with the development taking place in one and soon affecting the other.

On one hand there are advocates of the option that the 'new' conditions, the S (Aux) V(O) order, occurred first in main (matrix) clauses. They would later be generalised to

[66] McFadden's reference to Roberts (1997) has already been mentioned and commented on in section 3.2.
[67] As well as Bentz and Christiansen's proposal presented in 3.2.

coordinate and subordinate clauses (e.g. Traugott 1972: 160, Traugott 1992: 273, Hock 1986: 332). The main-to-subordinate clause direction of change would agree with the generally acknowledged tendency for dependent clauses to be more 'conservative'[68] than the main clauses in the process of syntactic change (e.g. Givón 1979b: 246, cf. Kohonen 1978: 133, Fischer 1992: 371).

On the other hand, there are proponents of the view that the future syntax of English became "grammaticised in subordinate clauses two hundred years before" it appeared in the main clause environment, as argued by Stockwell and Minkova (1991: 381, also Jucker 1990: 31-42 and more recently Allen 2000: 18-20, cf. Bean 1983: 109 and 137). According to Stockwell and Minkova, the order shift in subordinate clauses, appearing to some extent due to the surface manifestation produced by verb-raising, was initially created by analogy from the V2 order of main clauses. Following decliticisation of pronouns, visible in late fourteenth century texts, at a time when main clauses still had V2, in the subordinate clause environment it was no longer important to distinguish clitic pronouns from full NPs because both would precede the verb. The early prevalence of SV order in subordinate clauses was, according to Stockwell and Minkova, a key factor in the subsequent reanalysis of the main-clause order into V3 (SVX) (Stockwell and Minkova 1991: 400). In the same vein, Jucker (1990) asserts that matrix clauses could not lead the change to SV. The preponderance of the V2 pattern in these clauses is evidence against a reanalysis on the part of the child in the process of acquisition. By contrast, OV was disappearing quickly from the subordinate clauses, which put them ahead of the main clauses in word order changes (1990: 42). Lastly, Allen's account (2000) is especially worth mentioning because it provides a counter-argument to Lightfoot's claim (1991) that children during the process of language acquisition could only use the input from main clauses, which, if true, would remove subordinate clauses from the process of syntactic change. The robustness of verb-final order in early Middle English southern texts, according to Allen, does not support Lightfoot's "degree-0 learnability" (2000: 18-19). Interestingly, she also challenges Jucker's (1990) view of the dramatic decrease of OV order in subordinate clauses and provides evidence to show that the change was more gradual (2000: 20).

It is clear from the outline just presented that the status of main and subordinate clauses as being either conservative or innovative in syntactic change remains unsettled. One has to be cautious in relying strictly on the widely held view that subordinate clauses lag behind in the process of change, especially when dealing with such complex changes as word order shifts. Studies from other languages, for example by Schøsler (2001) on changes in French morphosyntax, show that subordinate clauses can be innovative in one respect (the loss of Pro-drop and stylistic inversion) and conservative in another (preservation of the two-case system and nonacceptance of V2) (2001:182).

The results of my analysis presented in chapter 6 below will shed more light on the complexities and preferences of word order patterns of the main clauses as contrasted with those found in subordinate clauses.

---

[68] "(…) in general subordinate clauses change their syntax more slowly" (Givón 1979b: 246).

4.4 FOREIGN IMPACT ON EARLY ENGLISH STRUCTURES

Along with the Scandinavian input, brought to the foreground in the previous sections, early English experienced influences from Celtic, Latin, its Romance successor, i.e. (Old) French and Dutch. Each of the listed external pressures had their own, substantial share in the shaping of *Vox Anglica* especially at the early stages of its development. The most visible effects of these foreign stimuli, quite expectedly, were (and still are) found in the English lexicon. Borrowings, by some estimates, constitute 75 per cent of the total of English vocabulary, with most of them coming from French and Latin (e.g. Thomason 2001: 10). There is also clear evidence of foreign interference present in the areas of morphology, both inflectional and derivational. Though much scantier, the influence on English phonology can be identified as well. Finally, owing to some recent studies, potential impact from external forces has also been identified in the domain that according to the implicational hierarchy proposed by Thomason (2001: 64) is known to be penetrated last in contact situations, viz. syntax.

The following sections cover predominantly those external influences that affected early English structures. Moreover, since the present study focuses on impact variables that are geographically conditioned, the outline of external influences will largely explore the Celtic and Scandinavian contribution to the English language. Consequently, the impact from other languages will be treated only briefly, as it involves much less geographical variation (e.g. Hogg and Denison 2006b: 15).

4.4.1 Celtic influence

The impact of (British) Celtic languages on Old English is recognised, along with the Latin influence, as one of the first instances of external pressure. The traditional view on the early Anglo-Celtic contact holds that the language of the Britons has provided almost no input to Old English or to its subsequent stages, apart from lexical instances referring to topography (river names, place names) and some personal names (e.g. Jespersen 1935: 34-5). However, recent studies on this earliest linguistic encounter have led to reassessments of the scope of the Celtic influence, especially with respect to English morphosyntax (and phonology). The revival of the topic began notably with the work of Poussa (1990), in which the emergence of the periphrastic *do* construction in English was attributed to contact with the British. The origins of periphrastic *do*, according to Poussa, are found in the early Celtic-English contact in the southwest of England. The development subsequently percolated to other areas. To support her claim, Poussa points to the first instances of the construction which appear in western texts of the thirteenth century. The eastern side received the feature a century later. The link between the parallel Celtic periphrastic constructions and the English usages which followed suit was formed, as Poussa puts it, in a process of creolisation, recognised as a contact phenomenon which, among others, has been responsible for generating an auxiliary-friendly syntactic environment, including periphrastic *do* (Poussa 1990: 407-434). However, as with many early proposals on the potential Celtic influence on the English syntax, Poussa's account has been viewed as speculative. The rise of DO was revitalised later by van der Auwera and Genee (2002), examining the question from an areal and typological standpoint. They see English as unique

among the Germanic language family with respect to the presence of the periphrastic *do* construction. The feature, according to van der Auvera and Genee, is strongest in the westernmost Germanic branch as well as in the Brythonic section of the Celtic family (Welsh and Breton), the one known to have been in direct contact with English for the longest period of time (van der Auwera and Genee 2002: 283-307). Similar to the views of Poussa, they emphasise the importance of the earliest occurrence of the construction in the SW dialects of ME (cf. Filppula 2010: 435)[69].

To continue with general typological influence of Celtic on English syntax and the issues around morpho-phonology, there are scholars such as Tristram (1999), who attributes the loss of inflections in late OE to contacts with British Celtic. Both English and Celtic have moved from synthetic to analytic means of marking syntactic functions. A powerful argument in favour of the Celtic input is that this trend, especially changes in noun declension[70], is attested earlier in Brythonic (Welsh) than in English (Tristram 1999, 2002a). Contact-induced changes in morphosyntax are also discussed in the work of Raymond Hickey (e.g. 1995). He considers the contact and convergence scenarios, discussing a number of syntactic phenomena, notably the Northern Subject Rule and the progressive[71] (also treated in Filppula 2010: 435). The former has also been discussed by Vennemann (2001), who mainly relies on the arguments presented by Klemola (2000) and Hamp (1975-1976) with respect to the Celtic impact on English and its dialects. According to Vennemann, the Northern Subject Rule is a substratum feature, having its source ultimately in the prehistoric Semitic substratum in Insular Celtic. From there it would have been subsequently passed on to English, with the evidence still existing in some traditional dialects of northern England and Scotland. In addition, Vennemann (2001) also discusses the verbal nouns (the *–ing* form of verbs) and the 'internal possessor construction'. In his opinion, the transparent formal and functional parallel between the English verbal noun/progressive constructions and its Welsh equivalents produces clear evidence for Celtic substratum influence on English. The Anglo-Saxon present participle ending in *–inde* or *–ande* was, in the course of time, replaced by the 'Celtic motivated' verbal noun construction with the *–ung/ing* suffix[72]. With respect to the 'internal possessor construction', too, parallels are found between English and Welsh (also Irish) but not German (Vennemann 2000: 399–406).

External influence of Celtic provenance also has to be considered when investigating the Old English (West Saxon) twofold paradigm of *to be*[73]. The paradigm, accord-

---

[69] For the geographic distribution of unstressed periphrastic DO in traditional dialects of present-day English, which to some extent mirrors the distribution of the feature during earlier historical stages with a Celtic connection, see Klemola (2002: 199-210).

[70] Excluding verbal morphology (Filppula *et al.* 2002: 13).

[71] For the origin of the English progressive – the nonstandard usages of the feature in Irish (as well as other 'Celtic Englishes') and its geographic distribution in traditional English dialects, see Filppula (2003).

[72] Vennemann (2011) upholds his view on the contact-induced development of the English progressive. Earlier studies which mention the possibility of Celtic impact on the English progressive include the work of Dal (1952) and Braaten (1967). Recently, Mittendorf and Poppe (2000), Poppe (2003), Filppula (2003) and Ahlqvist (2010) continue to support this claim.

[73] One paradigm begins in a vowel in the singular and with an *s-* in the plural (existential present). The other has forms beginning with *b-* (habitual present). They are both inherited from Indo-European (Cambpell 1959:349, Lutz 2009: 233).

ing to Hickey (2012), stands out in the Germanic context, as the other languages in the western branch have combined the two paradigms into one (2012: 500). However, one finds parallels to these paradigms in Brythonic, which have a similar syntax and semantics, for example the present-tense forms of *to be* in Middle Welsh. Lutz (2009) suggests[74] that the Old English double paradigm emerged through earlier transfer, as the Celtic speakers shifted to the language of the Germanic invaders who settled on the Isle (2009: 227, 238, 244). Schumacher's study (2007) also points to a clear dichotomy between the use of an unmarked *s*-stem and a marked *b*-stem for the verb to be (present indicative) in the Celtic languages (2007: 201-2). Still, Schumacher, unlike Lutz, is of the opinion that the Old English double paradigm has its source in much earlier contact of Celtic with West Germanic (North and East Germanic excluded) (2007: 193-202).

The latest studies into a number of structural characteristics in English, conducted by the group of researchers such as Filppula (Filppula *et al*. 2008, Filppula 2009), have revealed that the features in question can be seen to have emerged as substratal (or adstratal) contact influences from Celtic. Among them there are the already mentioned 'internal possessor construction', periphrastic-*do*, the *–ing* form of verbs as well as the *self*-forms of personal pronouns and *it*-clefting (Filppula 2009: 289). Last but not least, there is David White (2002, 2003) discussing the Celtic input to English from a typological perspective and in connection with the already mentioned simplification processes. He claims that the loss of case and grammatical gender in early English was triggered by the joint and converging impact from Norse (in the northern Britain) and from Brittonic (in the north and south-west). The same process was responsible, according to White, for the loss of inflectional variation in the definite article and adjectives (White 2002: 167). White maintains that it was the double dose of language contact which was responsible for the North being the most innovating area in the early English period. By contrast, the English southeast would be placed on the opposite side of the spectrum, receiving the least amount of external influence thus being the most conservative area. East Anglia and the southwest were set in the middle, as they went through one of the contact instances each (White 2002: 168, already mentioned in section 4.2). The proposals put forth by White will be considered again in section 4.4.2, where the main features of the Norse impact on English morphosyntax are discussed and in section 4.5.2, which addresses the question of the applicability of the wave model to the changes at issue.

4.4.2 Effects of contact with Old Norse

The Scandinavian legacy in the British Isles was (and for the most part still is) significant and lasting. The Viking period witnessed many important developments, especially with respect to urban and rural life. Expanding (founding) cities such as York and Lincoln along with the rest of the Five Boroughs (and overseas Dublin), building up trade networks with new markets opened, reinventing local administration down to the lowest ranks, implementing significant changes in structures of

[74] She follows Keller (1925) who was the first to claim that the two-fold paradigm was an Old English innovation under Celtic influence (Lutz 2009: 232).

estates (e.g. Roesdahl 1991: 56, Wainwright 1975: 214, Higham 1993: 260, Logan 1983: 138), these are just a handful of notable instances. However, by far the most impressive effects of Scandinavian influence are without doubt found in the realms of the English language. The most indelible mark has been left, expectedly, on English lexis, which extends beyond the repertoire of everyday, "need based" borrowing of terms designating new concepts (e.g. Dance 2012: 1727). Some Scandinavian impact has also been noted with respect to English phonology. Lastly and most importantly, quite a powerful influence has been exerted on the spheres of English morphosyntax. Many of the restructurings in that last area, the subject of the current study, contributed to profound changes which altered the make-up of English forever.

Although the current description is largely limited to the Scandinavian impact on early English morphosyntax, some remarks concerning borrowed lexis need to be made. The number and character of these loans shed light on the nature of the contact at issue. Out of an immense number of Scandinavian borrowings which have found their way into *Vox Anglica* over the centuries, approximately six hundred instances still flourish in present everyday English (Roesdahl 1991: 245). Combined with the hundreds of others preserved in rural dialects, the grand total of words of Scandinavian origin may easily reach two thousand, if not more (e.g. Geipel 1971: 70). The substantial number of words of Scandinavian origin, which, to an extent, superseded the early native vocabulary, has led some scholars to conclude that the contact was, indeed, responsible for significant relexification (e.g. Milroy 1984: 11). When it comes to the period of incorporation of borrowings, the first instances of Scandinavian items are noted already as early as the tenth century (Jespersen 1935: 58). However, the majority of lexical novelties do not occur in written records until the Middle English times, from the twelfth to the fifteenth century. Dance (2012) points out that it is highly likely for many ON forms, though recorded in texts at a much later stage, to have been regularly utilised in spoken English before the beginning of ME (2012: 1732). With respect to the quantity of borrowings specifying the character of the contact situation presently discussed, the small number of items attested in written data in the earlier period could be indicative of a cultural clash, as suggested by Kastovsky (2006). We could be dealing with a situation where the lexical novelties refer to items and concepts unknown to the receiving language (2006: 224)[75]. On the other hand, the flood of Scandinavian borrowings, including basic vocabulary and function words and along with subsequent incorporation of some of their morphosyntactic features during the ME period suggests language death with parallel language shift on the Scandinavian side. The North Germanic population was steadily switching to English with bilingualism withering at the same time (e.g Kastosky 2006: 224, Tristram 2002: 125, cf. Townend 2002: 201-207). As regards the relation between the borrowings and word classes, quite expectedly the members of the 'open' class (nouns, verbs, adjectives) represent the larger part of the borrowed material, with the items

[75] Otherwise, the scanty number of Scandinavian borrowings in OE could result from the simple lack of texts (Burnley 1992: 418). Next, the similarity (indistinguishability) of native words and Nordic forms, due to genetic closeness, might also be a relevant factor (Dance 2012: 1728-9). Finally, it could be the case of a lag between the actual contact instance and the act of assimilation of new forms (Millward 1996: 100).

belonging to the 'closed' class (pronouns[76], conjunctions, prepositions) constituting a minority (cf. Bynon 1977: 231, also Matras 2010: 78). If we take into account the fact that "the class membership of borrowed items will depend on the nature of the sociolinguistic context in which the borrowing takes place" (McMahon 1994: 208, also Matras 2010: 78), it becomes obvious why so many of the Scandinavian forms in English include content words denoting everyday items and activities. The two nations were very much alike, not only linguistically. They got used to each other quickly, both eagerly attending to daily chores, with the affected vocabulary reflecting this particular aspect of life. It will be shown in the following paragraphs that the existence of a common core vocabulary can have serious implications on the stability of the morphosyntactic systems of the languages in contact. What is more, the character of some borrowings will reveal why scholars distinguish different types of influence in the early Anglo-Scandinavian contact setting (section 4.5.1).

Place-names, regarded by Fellows-Jensen as "the best evidence for settlement by Scandinavian-speaking people in England in the Viking period" (Fellows-Jensen 1991: 337), deserve a special mention at this point. More than 1,400 place-names of Scandinavian origin have been counted in England. A substantial number of these are found in the areas of the medieval Danelaw, with some of its districts having as much as 75 per cent of the place nomenclature generated on a Scandinavian base (Hughes 2000: 95). When it comes to the manner in which foreign place-names influenced the native ones, Lindkvist (1912) lists a few possible routes of name modification. A Scandinavian name could be introduced in full (*Forss*) or the foreign part (affix) would be attached to a native form creating a hybrid (*Lundertorp*[77]). A place with a native root could often be renamed (*Northworthige* changed into *Deorabŷ* now *Derby*). Lastly, a native name could get partially 'Scandinavised' (*Eoforwīc* was transformed into *Jórvík* now *York*) (Lindkvist 1912: 51). Striking cases of place-name formation with a Scandinavian tint are those involving the (additional) use of ON morphological inflections. Some of the notable examples include *-ar/-a* (unreduced GEN sing.) *Help**er**thorpe* (*Hjalp's village*) and *–um* (DAT pl.) *Both**am*** (*búð**um*** i.e **at** the booths) (Geipel 1971: 123-4).

To continue with morphology, apart from the derivational novelties of Scandinavian origin mentioned when discussing place-names, there are instances such as ME *–laic*, *-lec* (cf. OIcel *–leikr* and contrast OE *–lāc*) used to create abstract nouns upon adjectives, e.g. *gōdlec* 'goodness' (Dance 2012: 1735). The impact of ON is also found in some verbal suffixes, e.g. *–n* (as in 'thinken') and *–l* (as in 'sparkle') (Miller 2012: 128). Moreover, new forms from ON also included elements belonging to inflectional morphology. Examples include the present participle ending *-and/-end/-ind* of late OE and early ME, superseded by *-ing* in later English (e.g. Fippula 2010: 436, cf. Miller 2012: 130-1). Scandinavian impact has also been recognised in the process of the generalisation of *–(e)s* genitives and plurals (discussion in Miller 2012: 132-134).

---

[76] Examples include the notable 3rd person pl. pronoun *they* along with *them*, *their* (cf. ON *þeir*, *þeim*, *þeirra*), which displaced OE equivalents: NOM/ACC *hie*, DAT *him*, GEN *hira/heora*, to potentially clear up the confusion with forms of the singular.

[77] The most commonly appearing affixes, apart from the aforementioned *–thorp*, include endings such as *–by* and *–toft*.

On the other side of the spectrum of morphological changes there is a process commonly identified in contact situations, the petering out of inflectional morphology. Longstanding interactions with Scandinavian population has been held responsible for the (acceleration of) syncretism and loss of the majority of inflectional forms (e.g. O'Neil 1978, Townend 2002). The commonly accepted scenario assumes that with the two related languages, with so many (lexical) parallels (e.g. Kastovsky 2006: 224, Mitchell and Robinson 1983: 132), the inflectional endings (especially case endings on NPs)[78] (e.g. Iglesias–Rábade 2003: 86) would at best be ornamental (e.g. White 2002: 157) and, at worst, they would put obstacles in the way of mutual understanding[79] (e.g. Mitchell 1994: 164, Curzan 2003: 52, cf. Holm 2010, 254). Regardless of how big the motivation to eradicate the endings was, they indeed differed between the two languages. When comparing only their nominal case systems, marked dissimilarity would be distinguished. The most prominent contrast lay in the presence of the ON *–(a)r* ending (NOM sing. and pl., ACC pl. and GEN sing.), which would be completely alien to the early English speakers. In the same manner, the English endings *–n* and *–e* (ACC and DAT, sing. and pl.) could be unrecognisable to the early Scandinavians. The ON DAT singular ending *–i,* too, was absent from the OE paradigm. Possible confusion could also arise with the *–u* ending, which in English was reserved for plural (NOM and ACC) whereas in ON it belonged to the singular (ACC and DAT). Some overlap could be found with the *–(e)s* and *–a* endings but only as far as the GEN environment. Finally, the only transparent parallel was identified for plural DAT *–um* (Gordon 1981: 283-289). Consequently, the speakers in contact could be faced with over 85 per cent of potential ambiguity on the nominal case surface realisation. Indeed, the confusion resulting from the clash of two distinct inflectional paradigms is sufficiently attested. The instances of Scandinavian place-names, listed above (incl. *Help**er**thorpe*), incorporated along with the ON inflections are definite proof of, at least, English ignorance of early Scandinavian inflectional grammar. Examples including the integration of the distinctive GEN sing. *–(a)r* ending are also found outside the realm of toponyms. The form is also preserved in Chaucer's *nightertale* (cf. ON *náttar þeli*) and the one designating masculine noun inflection found in the word *hagherlych* (Burnley 1992: 422).

Naturally, the Scandinavian impact on the fast pace of inflectional erosion was not prevalent exclusively in the noun morphology. It would affect the entire NP environment, including the marking of grammatical gender and that of the definite article (e.g. McWhorter 2002, Curzan 2003, also Dance 2012: 1735), with the domain extending sometimes to larger contexts, i.e. those of a sentence or text. What is more, studies by Kroch *et al.* (2000) presented earlier in section 4.3.1.1, show that the contact with early Scandinavian would entail the attrition of verbal endings[80] as well, with grave consequences for syntax (see also below). The inflections would be eroding at a much

[78] Studies show that in contact situations the loss of (noun) morphology is related to "the presence of a common core vocabulary and/or general intertranslability" (e.g. Danchev 1991: 119f).
[79] Considerable confusion of endings was already present on account of the fixing of the main (Germanic) stress. The inflectional contrast between the two languages, thriving in bilingual communities, certainly increased the volume of misunderstanding (e.g. Mitchell and Robinson 1983: 132).
[80] The inflectional erosion in the North occurred, according to Kroch *et al.*, due to imperfect learning of English by (adult) Scandinavians, who wanted to integrate themselves into the community (2000: 386).

greater tempo in the areas where Scandinavian presence was felt most – the North and parts of the East Midlands, which are the Danelaw areas. Textual evidence unequivocally supports that claim (e.g. Milroy 1992b: 182).

As for particular contact outcomes, it has been proposed that the longstanding interaction between early English and Scandinavian could have led to (a mild) creolisation, whereby the language in contact sheds its endings on account of the universal tendency for morphological simplification to occur in creole situations (e.g. Fisiak 1977, Poussa 1982, Milroy 1984, Danchev 1988 (1991), Ruiz Moneva 1997). In addition, there are scenarios of combined external input, where creolisation is a process which involves both the Scandinavian and French languages (Gerritsen 1984: 117, also Prins 1948, Workman 1972 quoted in Danchev 1991: 115, Bailey and Maroldt 1977). Tristram (2004), however, remarks that in cases of contact between languages which are sufficiently different (as OE and ON would be, according to some scholars excluding Tristram), in order to enable efficient communication, pidginisation would have to occur before creolisation (Tristram 2004: 94). Furthermore, she claims that the early English and Scandinavian languages were very similar to one another and the speakers of both could communicate with little effort. As regards the occurrence of pidgins before creoles, examples provided by Noonan (2010) clearly show that some creoles need not develop from pidgins (Noonan 2010: 60). Schiffman, too, speaks of creole languages (or languages with transparent creole traits) which have not gone through "the supposedly requisite stages of pidginisation" (Shiffman 2010: 745).

Crystal (2003) modifies the details of the Anglo-Scandinavian creolisation scenario by claiming that the contact situation between the two ethnic groups could have "led to the emergence of a pidgin-like variety, perhaps even eventually to a kind of creole which was used as a *lingua franca*" (Crystal 2003: 32). McWhorter (2002), too, points to the possibility of existence of 'pidgin English' among the first wave of Scandinavian speakers, "reflecting the limitations of adult language-learning capabilities" (McWhorter 2002: 261f). With respect to the correlation between the relatedness of the languages in contact and potential outcomes, indeed, the early Anglo-Scandinavian cross-encounters fit more the koinéization scenario than the one leading to the emergence of pidgins/creoles (c.f. Thomason and Kaufman 1988: 307). Following that logic, Dawson (2003) describes the contact situation in eastern and northern England as a koiné, the consequence of a close genetic relationship and (hence) typological proximity between the two languages (quoted in Roberge 2010: 422, cf. Noonan 2010: 57, Holm 2010: 254). Other scholars, on the other hand, propose inflectional loss within a contact situation which does not lead to any specific outcomes. Milroy (1992b), for example, claims that:

> *the advanced inflectional loss in twelfth- to thirteenth-century east midland dialects, (…), may be in some way associated with heavy Danish settlement in these areas – even if language varieties that resulted from this were not creoles (Milroy 1992b: 204).*

Milroy's account agrees with many currently held views that the contact between the early English and Norse populations did not need to create a new language variety

and that multilingualism was sufficient to account for the peculiar language situation in medieval England (e.g. Townend 2002: 196-201, Trotter 2012: 1791). The arguments supporting creolisation appear to be too strong, while recent studies are more moderate in presenting potential results of the early Anglo-Scandinavian language contact (Dance 2012: 1728).

Other morphosyntactic developments attributed to Scandinavian influence possibly include the ultimate selection of the Northern (and Midland) form of the verb *to be*: ME *are(n)* (OE – Mercian *earon*) (cf. ON 1 pl. *erum*, 2 pl. *eruð*, 3 pl. *eru*). The form overrode the popularity of the southern plural present indicative form (OE *sindon*, *sint*) (e.g. Roberge 2010: 421). Scandinavian impact has also been implicated in the development of the so-called contact clauses, whereby the relative pronoun is omitted in relative clauses. Similar constructions were found in ON (Jespersen 1935: 76, Dekeyser 1986: 114). Moreover, the increasing use of periphrastic perfect HAVE during ME seems to have been reinforced by parallel patterns in early Scandinavian. These can still be found in Modern Icelandic (*ég hef komið* 'I have come') and Swedish (*Vi har rest till Spanien förr* 'We have gone to Spain before') (McWhorter 2002: 236-8, 258). Contact with Norse has also been influential in the introduction of the phrasal verb type e.g. *come on*, *make up* (Poussa 1982: 73, also Dance 2012: 1735); otherwise the tendency for these types of verbs is claimed to have been strengthened due to Scandinavian parallels (Hiltunen 1983: 42-4). Rules for the use of *shall* and *will*, too, seem to have been applied in English on account of similar patterns in early Scandinavian (e.g. Jespersen 1935: 76).

Finally, there are accounts which implicate ON in changes in early English word order, with the altering of some of the aspects of V2 syntax and the general shift to VO being the notable instances. An important contribution dealing with the extent of Scandinavian influence on early English word order is represented by the work of Kroch, Taylor and Ringe (2000), mentioned already in section 4.3.1.1 and in the paragraph above. Recent investigations around the extent of Scandinavian impact on early English word order also include the work of Trips (2002). She aims to show ON impact on ME syntax by analysing various constructions in data from Early English. Her study primarily focuses on the *Ormulum* (2002: 3). Trips claims that the contact with the early Scandinavian language is responsible for the shift from OV to VO (Trips 2002: 331, 333). She tries to find features of Scandinavian syntax in the Early English data: object shift, scrambling, the V2 constraint and stylistic fronting. These instances of ON influence on early English syntax should, according to Trips, show that the OV-to-VO change can be attributed to this particular contact situation. Although providing numerous relevant points, her proposal has also met with some criticisms. One of the weaknesses mentioned is the choice of the text for her database, as remarked by Cloutier (2005). Findings based on this work of verse, which is "exceptional on many counts" according to Cloutier, cannot be used to make generalisations on the grammatical status quo of early ME and much less on the properties of the spoken language of the period in question[81] (2005: 181-2). Others have pointed to different rea-

[81] The present study is based on the prose data precisely due to its proximity to the spoken word and the lack of stylistic intricacies that characterise poetry (see section 5.3.2 for details).

sons behind the popularity of the features analysed by Trips. For example, a study by Masayuki (2010) reveals that the effect of stylistic fronting is not due to Scandinavian influence but a shared property of ME in general (2010: 115).

The present study offers yet another take on how contact with ON could reinforce tendencies for rigid (S)VO. Here, the discussion on correlation between (the lack of) word order freedom and presence (or absence) of case is re-examined and emphasised. Interaction between the loss of (case) of inflections and preference for this particular syntactic ordering is commonly identified in language contact settings. Even though both processes could occur due to purely internal reshuffles, it has been shown that the connection between the two will be much more intense when external pressures are an additional yet relevant factor and when the two processes operate in conjunction, which seems to have occurred in English (Danchev 1991: 116 *passim*). Before the essentials on this correlation in connection to Anglo-Scandinavian contact instance are presented in greater detail (section 4.5), a brief account on the impact from other languages on early English structures is needed.

4.4.3 Influence from other languages

The influence of Latin on early English came in three waves. One arrived from the Continent, with contacts taking place before the Germanic tribes left their homelands. The other occurred through the influence from British Latin, as *Vox Latina* was, ultimately, the only written language in early post-Roman Britain (e.g. Schrijver 2002: 89). The last impact came with the Christianising of Britain, beginning with the coming of the missionaries sent from Rome in 597 (Townend 2006: 65). Of the three, the last one in particular left an enduring imprint onto the English language.

While the Latin impact on early English lexis is quite transparent and well recognised by scholars (e.g. Williams 1975: 57, Burnley 1992: 432-3, Millward 1996: 101, Townend 2006: 72, Kastovsky 2006: 250), the influence within the spheres of syntax has been largely ignored in studies, as remarked by Vezzosi (2012: 1716). Still, one can clearly see how Old English prose is derived from Latin originals. One of the features that has been proposed to have its origins in Old English texts containing interlinear glosses or translations from Latin originals is the English progressive (Mossé 1938). A factor speaking in favour of Latin influence is the relative infrequency of the construction in the poetical texts of the period. However, scholars such as Nickel (1966) are less enthusiastic about the impact of Latin on the development of the English progressive. His analysis of typical contexts in which the construction appears in OE texts revealed that the content clearly influenced by Latin structures did not always fully conform to the model of the corresponding Latin forms. Another feature the development of which is claimed to have been influenced by Latin is the cleft construction. Yet, studies by Ball (1991) and most recently by Filppula (2009) show that the impact of the classical language in these spheres of early English structures is not entirely plausible. According to Ball, OE authors sometimes seemed to avoid cleft constructions in their translations of Latin works (Ball 1991: 52). In addition, Filppula's study

of the Old English prose corpus (YCOE[82]) has shown that the texts in which clefting appeared were not based on Latin originals (Filppula 2009: 285).

The impact of French comes through the introduction of the language of the Normans after the marked date of 1066. The Romance language became the spoken[83] medium of the ruling class in England after the Conquest. Although English was used by the lower orders of society, the exquisite old written tradition subsided, with only a few records available for the two centuries after 1150 (Knowles 1997: 46). French continued to be used for some three hundred years until it began to be gradually superseded by English from around the mid-fourteenth century (Knowles 1997: 48). Indeed, Middle English appears to be a blend of OF and OE. However, with respect to semantics the "content owes a great deal to the despised[84] and unrecognised *faus franceis d'Angleterre*, i.e. 'the faulty French of England" (Diensberg 1996: 259).

As in the case of Latin, the effects of the French influence are found predominantly in the English lexicon. Nonetheless, there were signs of this foreign influence in early English in the domain of morphosyntax as well. The items were either introduced afresh or popularised by their French counterparts. The French impact was present in some uses of prepositions such as *at*, *by* or *in* (e.g. Filppula 2010: 437). In addition, there was the introduction of new prepositional options (i.e. reanalysis of participles as prepositions) which could be based on French usages[85]. Among the notable examples are *according to*, *during*[86] (Molencki 2011: 11-2, referred to as quasi-prepositions.) Some of the adverbial uses are claimed to have had their source in the French language introduced into medieval England, e.g. *albeit*, *as, very* (Filppula 2010: 437). Next, the use of *for* with infinitives in Middle English, as in one of the lines from Chaucer's Opening Prologue to *Canterbury Tales*: *And wente for to doon his pilgrimage*, could have been calqued from French (Miller 2002: 187-241, 2012: 185). Next, the internal pressures oriented around the gradual shift from OV to VO combined with the borrowing from French both had an effect on the periphrastic genitive construction (Lehmann 1973: 185-6). In early Middle English, the OE manner of genitives preceding nouns was still maintained, even within phrases which had no overt genitive marker. However, during late ME the genitive began to be visibly placed after its noun. Already around the fourteenth century about 85 per cent of genitive constructions consisted of a postposed genitival phrase with *of* (by contrast only 1 per cent contributed to the total in the 10th c.) (also in Miller 2012: 186, overview in Szmrecsanyi 2013: 61, section 4.2). Another construction which was made more prevalent due to French influence was the Middle English partitive, e.g. *Hij ne eten of oxe ne of swyne* 'they eat neither ox nor swine' (Miller 2002: 333 f20, recapitulated 2012: 186). Further, French influence came to be associated with the rise of the progressive and the *it*-cleft construction in early English. There are parallels with respect to these forms found in both languages (Filppula 2010: 437, 440, 441-3).

French impact has also been implicated in more complex and overarching devel-

---

[82] *The York-Toronto Corpus of Old English Prose*.
[83] With Latin dominating in the written realm.
[84] Both English and French were seen as inferior to Latin in the early medieval period (Knowles 1997: 47).
[85] Latin influence could be instrumental here as well, esp. with *providing (that)* (Molencki 2011: 13)
[86] Occasionally as a postposition (Molencki 2011: 12-13).

opments such as word order changes. Haeberli (2010), for example, sees the beginning of the loss of the V2 order as initiated through contact with French (2010: 143-163). His study adds an interesting angle to the complex situation around the operation of the V2 constraint in early English – the southern versus the northern (Scandinavian) V2 type (covered in section 4.3.1.1). Further, views according to which the emergence of SVO in (early) English was induced by contact with French have also been presented (e.g. Weinstock 2006; Ingham 2009). Miller (2012), however, maintains that the French impact on SVO could be problematic. One of the reasons is connected to the fact that during the final stage influence from early Scandinavian has to be considered. The possibility of shared innovation, according to Miller, does not necessarily have to apply either (2012: 185). Lastly, when it comes to the types of outcomes generated by French influence on medieval English, scenarios of creolisation or language mixture were offered as early as 1977 by Bailey and Maroldt. They saw the Middle English language as a creole, created out of the English and French languages in the period after the 11th century. The peculiar linguistic situation would additionally be enhanced by the earlier conditions prevailing in the North. The time between the end of the 10th and mid 11th centuries witnessed the arrival of a creole which arose in the North as a result of contact between Old English and Old Norse (Thomason and Kaufman 1988: 309-10). However, as pointed out earlier in sections 3.1 and 4.4.2, recent studies indicate that creolisation phenomena with reference to Middle English do not seem to apply.

Some attention needs to be dedicated to potential (structural) influence from Dutch. Traditional accounts on immigrant populations to the Isles within the period of the eleventh to thirteenth centuries do not explicitly mention groups from the Low Countries. According to Bense (1925) "[t]he term 'Norman' has been used to indicate the miscellaneous host of […] men which William of Normandy led to the conquest of England […]." (Bense 1925: 7). Hendricks (2012) picks up on Bense's observation and states that relying on the term 'Norman' indeed obscures the complexity of the contact situation in the British Isles at that time. Thus, within the Norman army one could easily encounter Flemings who were paid for their services in land, given estates and plots throughout the country[87]. The subsequent influx of people from the Continent, too, included people from Flanders (Hendricks 2012: 1663-4). As for the impact on early English structures, the discussion, though much scantier than instances of Low Dutch lexical borrowings, includes considerations around the *th*-stopping in which /d/ replaced /ð/. Examples such as *de* 'the' and *dis* 'this' are found in numerous Middle English texts, as researched by Samuels (1971: 11). According to him, the process occurred in all contexts by the fifteenth century in data from the Kentish dialect (spoken in Kent, East Sussex and East Surrey). Wakelin (1977), too, sees a possible Low Dutch influence in early English in the closure of /ð/ to /d/ in south-west dialects (1977: 92). The potential later impact includes the loss of the 3rd person singular –s inflection in the East Anglian dialects, which, according to surviving evidence, switches from the

[87] Thomason and Kaufman (1988) assert that the Low Dutch speakers were not evenly spread over England. Instead, they focused on places where their interests lay. Thus, merchants and brewers would choose towns and cities. Dike-builders, on the other hand, would settle near the fenlands (1988: 322).

Middle English (eleventh to fifteenth century) *–th* form to zero. By the sixteenth century, the zero form becomes the norm, as suggested by Trudgill (2002: 185). The period in question parallels with the time of large-scale migration from the Low Countries and France into East Anglia. Finally, the claims of pronominal transfer (the enclitic/unstressed object form of 'she' and 'they' – *(h)is* or *(h)es*, occasionally spelled *hise*) from Flemish to Middle English dialects (coastal areas within East Midlands and the South) put forward by Thomason and Kaufman (1988: 323-25) have been questioned, as remarked by Hendricks (2012: 1668).

Sections 4.4.1 – 3 covered a wide array of morphosyntactic alterations in early English which had been reinforced through contacts with the many foreign populations settled on the English soil from the earliest days to the time of the last conquests. The majority of the accounts of external impact on English structures raise an immediate problem that quite a few of these changes, word order shifts included, are attributed to more than one contact scenario. If a particular contact instance is to be preferred over the others (in the present study – Scandinavian influence), it is essential to narrow the range of possibilities. One of the means to do it is to rely on dialect geography. The following sections offer some more input to that type of account.

4.5 MORPHOSYNTACTIC CHANGE IN EARLY ENGLISH AS CONTACT-INDUCED

So far, several facts about the emergence of the rigid SVO order in early English have been established. Firstly, the change was initiated already during the early stages of the development of the language. Within the Germanic family, English is the only one to generate this syntactic order without the underlying, typically Germanic, V2 rule. Secondly, the emergence of SVO was a complex process, arrived at as a result of amalgamation of secondary developments, occurring in the course of several centuries. Many of these developments were reinforced by the erosion of inflectional endings. Thirdly, the input of inflectional erosion in the emergence of SVO is traditionally regarded as a part of typological drift reinforced by phonological reduction. However, since English diverged so much from the family matrix and since most of the traditional approaches cannot account for the changes that occurred in English (e.g. Bean 1983: 39), other explanations have to be sought. Fourthly, a substantial portion of the motivation to turn to SVO can be found in the working of external pressures on English – Latin, Celtic, (Old) French and (early) Scandinavian. Of these, the Scandinavian input is known to have been implicated in changes of both the NP and VP environments, largely mediated through the simplification of inflectional morphology.

Various outcomes within the early Anglo-Scandinavian language contact have been proposed as instrumental in the morphosyntactic change - pidginisation/creolisation phenomena as well as those associated with koinéisation. There are also options available which would not lean on any of the outcomes just mentioned. This last possibility is especially appealing considering the fact that contact results such as creolisation are not considerably divergent from those generated by normal language

change (Schiffman 2010: 744). The lines of development distinguishing one outcome from the other(s) are often blurred, frequently with differences in quantity rather than quality (e.g. Schiffman 2010: 744). However, there are a few aspects of the Anglo-Scandinavian contact which remain constant. The linguistic encounters in question were longstanding and intense. The languages in contact belonged to the same language family (hence their typological proximity). The contact situation was one that could easily lead to bilingualism (e.g. Mitchell and Robinson 1983: 132), and with the speakers of both populations willing to facilitate communication, some adjustments and merging of the two linguistic systems could take place, generating convergence.

4.5.1 Germanic languages in contact: convergence

Convergence, defined as the increasing agreement of languages not only in terms of vocabulary but also with respect to features of the overall structure of both (Hock 1986: 492; cf. Matras 2010: 68), is one of the types of contact-induced linguistic change. Two conditions need to be observed for convergence to take place. First, it arises only in the presence of extensive and long-standing bilingualism. Second, the participating languages need to be perceived as socially equal. Convergence will not come about if the languages in contact are in a kind of relationship that Hock describes as "a strong superstratum or substratum which would quickly oust the other(s)" (Hock 1986: 491). Further, it is vital to note that the process in question has an impact more on the syntax and morphology than on the vocabulary. Finally, the motivation behind convergence lies in the need for ease of learning and communicative efficiency. Thus, we are dealing with a community whose speakers need to use both languages (grammars) in order to communicate. The rules of the two languages would be easier to master if the grammars were more similar (McMahon 1994: 213).

According to Hock (1986: 492), prolonged bilingualism has important consequences which go beyond a mere long passing of time. The interaction of two (or more) languages leads to the creation of increasingly complex and 'mixed' interlanguages. With time, an array of those mixed varieties comes to exist alongside the earlier, less complex ones. As a result, it might be very difficult to establish which of the shared features originated where (cf. McMahon 1994: 214). The most significant observation, however, concerns the breadth of the scope of bilingualism and the aftermath of convergence. Hock claims that "it is not necessary for all speakers of the languages involved to be bilingual, or for all dialectal areas of these languages to be bilingual" (Hock 1986: 493). The outcomes of convergence can diffuse to new speakers or new dialect areas via the same mechanisms which allow other linguistic innovations to disperse.

The result of the language contact between the early English and Scandinavian populations fits the convergence scenario, as remarked by Görlach (1997: 23). He points to the formation of a new speech community among the speakers of Anglian Old English and Scandinavian dialects in the tenth and eleventh centuries. The community came into existence by population movement, the process motivated by invasion, subsequent battles and final settlement, as well as by the setting of new political boundaries, which then led to the emergence of territories under direct Scandinavian control, the Danelaw. The speakers of both ethnic groups had to com-

municate by means of their related languages, which led to a convergence of the two systems (Görlach 1997: 23). Since the communication between the Danes and the native English was of utmost importance, it would be highly advantageous for the Scandinavians settled in Britain to have become bilingual as soon as possible (Geipel 1971: 57). Fellows-Jensen (1991) notes that the Scandinavians had some knowledge of English as manifested by English place-names in Icelandic skaldic poetry (1991: 340). At the same time, the English would strive to understand people from the Danelaw. The extensive borrowing of Norse lexis, which often resulted in complete replacement of their English equivalents (e.g. Millward 1996: 106-7), is a patent indication of an acquisition process also on the English side. Still, as remarked by McWhorter (2002), "the impulse towards bilingualism would have been much stronger among the Vikings than among the English" (McWhorter 2002: 260, cf. Geipel 1971: 61-2).

The two languages were intelligible[88] to some extent from the very beginning and a mixed dialect may soon have developed in the Danelaw (cf. Roesdahl 1991: 245). Logan (1983) speaks of *Donsk tunga*[89] being used during the bilingual stage after which it became conflated with Old English to such a degree "that the term 'Anglo-Scandinavian' can be used to describe the language" (Logan 1983: 170). The confirmation for linguistic mixture to have taken place can be obtained from numerous Norse runic inscriptions, many of which were "naturalised as to be used by speakers and readers of English" (Wakelin 1988: 72). A similar tendency of mixing native and foreign forms occurred side-by-side in the Anglo-Saxon runes (e.g. Bugge 1921: 203). Other, non-linguistic input available which supports the claim, e.g. data retrieved from archaeology, is realised by Anglo-Scandinavian ornaments on crosses and hogbacks in Yorkshire (e.g. Roesdahl 1991: 249) along with other specimens of stone-carving, the tradition of which was not destroyed but revitalised by the early Scandinavians (Morris 1981: 223-44). Another linguistic fact testifying to the mixture of Anglian and Norse comes from hybrid place-names, with the so-called 'Grimston hybrids' being notable examples (e.g. Geipel 1971: 122, Wainwright 1975: 290, Wormald 1982: 135, Logan 1983: 168).

The extensive borrowing of words for everyday usage (e.g. Thorson 1936: 7, McWhorter 2002: 253) along with grammatical words, including the already mentioned peculiar instance of third person plural pronoun borrowing (e.g. Townend 2002: 205) present shining examples of unusually close and long-standing intermingling between the two ethnic groups (also Roberge 2010: 420). The latter is especially indicative of intense long-term contact and proficient bilingualism (e.g. Campbell 1997, Thomason and Everett 2005). The heavy borrowing of content words relating mainly to everyday vocabulary provides an additional input to a rather special relationship between the two cultures. Wilson (1976) remarks on the Anglo-Scandinavian interactions in the north and east of England:

> *One can trace quite clearly in the art of the north-east the successful blend of two different artistic traditions, a blend which illustrates the congruity of two different*

[88] There is ample textual evidence that supports mutual intelligibility (e.g. Townend 2002: 186).
[89] 'Danish tongue' – the expression used to collectively refer to the Scandinavian languages in the Viking Age and for a couple of centuries afterwards (e.g. Roesdahl 1991: 46, cf. Bibre 2001: 89 'Norse tongue').

> *artistic traditions, English and Scandinavian. The mixture illustrates the similarity of outlook of two different people, talking closely-related languages, living in the same area, worshipping the same god and having a certain continuity of political and even ecclesiastical organisation (Wilson 1976: 399).*

It is clear that the two nations were culturally equal, a fact acknowledged by many scholars, notably Laing and Laing (1979), who speak of the "cultural fusion of Angle and Dane in the north" (1979: 185). Townend (2002), in order to show that the two cultures were on a par with each other (i.e. two languages being "roughly adstratal"), considers the alternative scenarios on the relationship between Old English and Old Norse:

> *(...) if Old English were of much greater prestige one would expect the rapid death of Old Norse and few Norse loans in Old English; and if Old Norse were of much greater prestige one would expect many loans in Old English of a non-need nature, and certainly not the death of Old Norse (Townend 2002: 204).*

Trudgill (2010), amongst others, also believes that the relations between the two languages were maintained on the adstrate level (Trudgill 2010: 35, Hogg and Denison 2006b: 13, cf. Thomason and Kaufman 1988: 303). One can encounter, nonetheless, proposals suggesting both a Norse superstrate (e.g. Lutz 2012: 508-517, 2013: 562-590) and substrate effects from Norse (e.g. Miller 2012: 98). Perhaps the most apt observation on that account is provided by Dance (2012) who asserts that with such a complex contact situation all kinds of contact interactions could have existed at some point (2012: 1727).

As regards the areas of language affected by the contact, Thomason and Kaufman admit that the influence of Norse on English was pervasive, since the evidence of that influence can be found in all parts of the language. However, the impact, according to them, was not deep except in the lexicon. In addition, they claim that the early Scandinavian languages could not have changed the basic typology of Old English because the two languages were too much alike (Thomason and Kaufman 1988: 264). In contrast to Thomason and Kaufman, McWhorter (2002) demonstrates that the impact of Norse was indeed profound and what Thomason and Kaufman presented as the 'Norsification package' (Thomason and Kaufman 1988: 292-8) constitutes only a part of the relevant evidence (McWhorter 2002: 256-7). The Scandinavian impact on early English goes far beyond mere sound changes and lexical borrowings. Commenting on the close proximity of the two languages, McWhorter, relying on the example of the Riau dialect of Indonesian (Malay), seeks to show that even with closely related languages "reduction can play as significant a part in the outcome as exchange of materials" (McWhorter 2002: 261). He goes on to argue that when linguistic accommodation takes place, speakers commonly take part in the process by creating "a less overspecified and complexified" version of their language (ibid., 261-2).

Despite the similarities between OE and ON, one should keep in mind that the two languages, though related, maintained the status of being separate languages for a long time (Poussa 1982: 72, Roesdahl 1991: 46, A. and R. McMahon 2003: 7-55,

McWhorter 2002: f18)[90]. Besides, as McWhorter (2002) points out, a close genetic and typological proximity of the two languages does not hinder the extent to which modifications can be made on account of contact (2002: 230). As long as there is still a slot, an 'open space' available for a change (cf. Visser 163: 135) and if the speakers have the need to increase the ease of communication, changes will be introduced. To illustrate, in the areas of language already attested in many respects as almost identical, such as Old English and Norse lexis there are intriguing cases of merging of forms which occurred as a consequence of the contact between the two languages. Lexical borrowing from Old Norse involves instances of the fusion of form and meaning – a development reinforced by the phonological resemblance of OE and ON lexical forms and, more often than not, operating on account of their partially shared semantic spaces (Williams 1975: 193). Among the notable examples are the words *dream* (Serjeantson 1935: 74) and *dwell* (Jespersen 1935: 64-5), both of which had their original meaning shifted, i.e. replaced with that offered by the similar sounding ON forms of *draumr* and *dvelja*.

Apart from examples of lexico-semantic fusion, phonological similarity could also have operated in the process of smoothing contrasts within the article system. The replacement of *se* (masc.), *seo* (fem.) with forms containing the initial consonant */θ/* in early English, as mapped on ON *þœn* (masc.) and *þe* (fem.) are notable instances. Both languages already had the identical sounding/looking form in the neuter *þœt* (McWhorter 2002: 229, quoting Gordon 1981: 302). Further, in the area of morphology, the traditionally recognised shedding of (different sounding) inflections to facilitate communication between the two ethnic groups (e.g. O'Neil 1978: 256-260) serves as a neat instance of merging of the two language systems. Finally, there is the loss of clitic status on pronouns in early English (section 4.3.1.1), which occurred initially in the Scandinavian influenced areas (Fischer and van der Wurff 2006: 185). As examples of convergence within syntax one could mention a stronger leaning towards VO (regardless of changes in morphology) (e.g. Fischer et al. 2000: 175, Faarlund 2004: 161).

4.5.2 The wave theory as a model for explaining the changes in early English morphosyntax

An obvious difficulty in explaining the emergence of far-reaching morphosyntactic rearrangements in early English springs from the complexity of the linguistic change in question. There is a long distance from sound changes, bringing about morphological simplification to a wholesale syntactic shift, with a myriad of small but important alterations in between. Some of these changes had its starting place within the linguistic structures. Of these, a few were subsequently propelled by pressures from outside the language. There were also changes which found their origin in external pressures alone. The current study attempts to establish the extent of impact of foreign, especially Scandinavian interference on the emerging new syntax. Many of the rearrangements which led to emergence and prevalence of new word order, including the growing preference for VO, the loss of the validity of the V2 constraint, and the changing behaviour of personal pronouns, had at its base the loss of (case) inflec-

[90] A number of scholars suggest that early English and Scandinavian should be treated as dialects of the same language (e.g. Lass 1987: 52, Kastovsky 1992: 328-329, Townend 2002: 41, 60, 180, Davis 2006: 154).

tions. Therefore, it stands to reason to propose that the emergence of SVO would have been mediated through this process, which was already underway before the eve of Scandinavian incursions, settlement and assimilation, but which was later considerably augmented in this contact setting. To demonstrate a connection between the two processes and to eliminate other foreign influences potentially implicated in the shift, analysis of the dialect data has to be able to reveal a geographical nexus.

Much of the current research attempts to explain the Scandinavian influence on early English syntax through convergence – a process of elimination of differences between two related but linguistically independent languages, which at some point in the early history were geographically contiguous and whose territories became consolidated under the impact of "some political force with a single administrative and cultural centre" (Bynon 1977: 193). Analysis of the written records clearly shows that morphological simplification occurred at a specific point in time (which was late OE) and was initiated in a precise location, viz. the northernmost dialects of early English. The distribution of SVO in the data would have to replicate this layout, though naturally, at a later stage (ME). In order to follow the path of the emergence of the syntactic feature, the level of regional grammatical variation in the distribution of SVO is examined in this study both diachronically and synchronically, by means of the data extracted from parsed corpora of Old and Middle English. The model of linguistic change chosen to highlight the importance of the foreign impact in the morphosyntactic alterations at issue is Schmidt's wave theory, introduced in section 3.4.

The preference for SVO is measured at earlier (OE) and later (ME) stages of development of English, with clearly marked dialect information serving as a substitute for isogloss lines. Dialect boundaries can, at some point, correlate with successive stages of change both synchronically and diachronically, as no two dialects are ever identical with their preference for a feature (cf. Wolfram and Shilling-Estes 2003: 722 - transitional area). With the help of statistical tools, the level of dispersion can then be compared even within the same dialect where subperiodisation of the available material dissects a linguistic region into shorter time-span units. Eventually, potential focal areas - dialect sectors which exhibit the highest and most regular preference for the feature as well as peripheries - the units where the prevalence of SVO is less pronounced can be distinguished.

Since the content of the databases used in the current research provides only some idea of the characteristics of the spoken language in the historical period, the model used to explain the changes here discussed need not include many of the aspects affecting dispersion of the feature through socio-geographic space (cf. Biber *et al.* 1998: 252). The corpora in question comprise prose manuscripts, which already introduce an additional portion of text(type)-oriented parameters that ought to be addressed. The wave theory is therefore effective enough to point to parallels with respect to (sub) period and location. The repetitiveness of patterns within all studied perspectives, whether relating specifically to text types or particular grammatical environment, as filtered through time and space, should provide a substantial basis for making generalisations on the morphosyntactic climate of early English.

Finally, the phases of language change which the wave model highlights, i.e. the point of origin and the subsequent spread of the new feature through socio-geograph-

ic space, are found in the manner that the Scandinavian influence manifested itself on a dialectal map of early English, as remarked by Hogg and Denison (2006). They identify linguistic variation in Britain as one of the features which, indeed, separates the Scandinavian impact from other external influences. According to Hogg and Denison, the influence in question was a two-stage process. First, we have the Danelaw area, where the Scandinavian presence was originally felt most. It is the location of original contact between the two languages, which introduced Scandinavian features into the English of that district. Cumberland, Westmoreland, Yorkshire and parts of Lincolnshire, places where many Scandinavian loans still survive, all belong to this focal area (Samuels 1985: 272–4). Secondly, there is the successive spread of these features within English "by means of inter-dialectal contact" (Hogg and Denison 2006b: 15). By contrast, the French impact on English is more connected with differences between types of social language rather than geography. There is much less dialectal variation, and the focus is put on register. The relevant boundaries are not those of linguistic region but of the level of formality of particular texts. Influence from Latin follows a similar logic. The only contact instance which could be comparable with the nature of the early Scandinavian contacts is the impact from the Celtic languages. Just as the dialectal map of early medieval England may be divided into Scandinavian and non-Scandinavian sectors, the Celtic and non-Celtic zones could be identified too, as noted by White (2002). Showing the adequacy of one contact scenario over the other may prove challenging, considering that both Scandinavian and Celtic pressures are mentioned as potentially instrumental in the morphosyntactic change presently discussed and that the geographic areas of influence of both overlap to some extent, viz. the North (White 2002: 154). However, there are dialects which belong specifically to the sphere of one of the candidates but not the other – the East and the West Midlands. The former falls exclusively under Scandinavian control, while the latter is under influence from Celtic. It is important to keep this division in mind, when observing the distribution of the relevant features through dialectal space.

5 Aims, Methods and Data

5.1 RESEARCH QUESTIONS, AIMS AND HYPOTHESES

This study investigates one of the most important syntactic developments in English, viz. the emergence of the rigid SVO order. Beginnings of the process can be seen already in Old English, with the peak of the change occurring sometime during the Middle English period. Among the series of structural reshuffles, there was inflectional erosion and the collapse of the case system in particular, which has been regarded by some scholars (e.g. Robinson 1992: 166; also Bentz and Christiansen 2013: 45-61) as the ultimate reason for the subsequent growing preference for a rigid SVO order. Although both internal and external pressures have been identified as influential in the morphological simplification as well as the emergence of the new syntax that followed, this research focuses on external influence, the language contact situation that the early English had with the Scandinavian population from roughly the end of the eighth century onwards. The data used in this study have been extracted[91] from two parsed corpora of early English: *the York-Toronto-Helsinki Parsed Corpus of Old English Prose* (YCOE), and *the Penn-Helsinki Parsed Corpus of Middle English, second edition* (PPCME2).

The first research question addressed in this study is whether the word order change could be seen as a follow-up process to the inflectional erosion. More precisely, whether the loss of nominal case marking created a situation structurally unstable enough to reinforce a shift in the order of syntactic components. Establishment of a strict SVO order would be seen, at least to some extent, as a compensatory measure to solve potential ambiguities with respect to the marking of core grammatical functions, that of the subject and the object. Semantically, we are looking at an attempt to clarify the identification of Agent-Patient relationships. The correlation between presence/absence of case and the preference for SOV/SVO, respectively, has been attested in several previous studies, as shown in section 3.2. Greenberg's universal 41 (1963) already speaks of a trade-off between case marking and these word order systems. Recent accounts, including the proposal by Bentz and Christiansen (2013), who use data from WALS, confirm it (2013: 56-7). With respect to early English in particular, the SVO order was already a popular syntactic tendency, especially for configurations where subjects and objects were nouns (not pronouns) (e.g. Bean 1983: 139). Furthermore, the outline in section 4.3 pointed, among others, to scenarios which, indeed, emphasised the connection between the loss of cases on NPs and the growing preference for VO. That section also described accounts pointing to morphological erosion, case loss included, which was instrumental in the process of decliticisation of pronouns and the increased overtness of the subject. Both of these supported the

[91] By means of the Corpus Search programme (CSearch) [http://corpussearch.sourceforge.net/].

SVO layout. In this respect, the present research attempts to reconfirm the correlation mentioned above and to find out whether there is a convergence point on the dialectal map of the corpora material used here between the morphological simplification and the emergent (S)VO word order. The point in question would have to be located in the northernmost dialect sectors. Evidence from the written records created during OE and ME times show that this is where inflectional erosion began to transpire most, as described in section 4.2. In order to identify the convergence point as the place from which the new syntactic tendencies unfolded, one would have to distinguish similar repetitive patterns of (S)VO along with a visible contrast between the most innovative sectors and the remaining dialect sets.

As pointed out in sections 3.1 and 3.3.2, morphological simplification and preference for a fixed SVO word order are also common phenomena occurring in language contact situations. The second research question deals with the extent of the role of external pressures of such sort. More specifically, the current study explores the degree of impact of the early Anglo-Scandinavian contact situation on the morphosyntactic change at issue here, especially in the kind of quick-pace erosion of (case) inflections that we are dealing with here (e.g. Townend 2002: 201, Bentz and Christiansen 2013: 55-56). Indeed, a number of scholars have pointed to the attrition of cases occurring specifically as a result of this instance of language contact (section 4.4.2). The interaction of the Anglian and Scandinavian populations or rather the effects of that interaction coincide with the onset of the structural shift, revealed in the records of the period. Apart from the Scandinavian contribution to the change within the realisation of nominal subjects and objects, the previous sections (4.3.1.1 and 4.5.1) have also listed the early North Germanic input to the status of English pronouns as well as to the overall increase in VO structures. All of the listed tendencies, morphological simplification included, can be regarded as examples of the language contact scenario advocated in this research, viz. convergence. This study, therefore, aims to assess whether the convergence point correlates with the areas where Scandinavian influence was felt most. Dialect-wise, we are looking for, at least, the highest preference for VO and a high number of the entire SVO sequence tokens in the North along with the East Midlands sectors. On a more general level, this research examines how important external pressures are for the development of a language and how far-reaching the consequences of these pressures are. Therefore, an attempt will be made to deem the contact setting currently discussed as either a mere facilitator or as an effective disseminator of the changes investigated (cf. Matras 2010: 72).

The third research question focuses on the outcomes of convergence and the theoretical model chosen to test them, i.e. the wave theory. The present study aims to assess whether this convergence model could be used to account for the emergence of SVO, occurring due to, at least partially, externally motivated morphological simplification. As described in section 3.4, this theoretical framework stresses geographical rather than genetic factors. It has often been used to accompany or to displace a more traditional and abstract family tree model (e.g. McMahon 1994: 229). According to the wave theory, the level of spatial proximity matters more than the relatedness of the languages (or dialects) in contact. The theory also emphasises the social aspect, whereby an innovation spreads from one population to another, as speakers have

contacts with more than one community. With respect to the data analysis, this study endeavours to distinguish the pattern of change advocated by the wave model, i.e. the clearly marked focal area, encompassing the dialect sectors where the distribution of the new feature is the most prominent and most regular, as opposed to the peripheries, where the dispersal of the innovation is less pronounced. Overall, this research aims to revisit and reassess the explanatory power of the wave model both when it comes to its applicability in longitudinal comparative studies as well as in accounting for the changes that occur at the earlier stages of the development of a language (e.g. Renfrew 1989: 112, cf. Rankin 2003: 186, Wolfram and Schilling-Estes 2003: 721-32).

Finally, the fourth research question tackles the issue of the usefulness of parsed diachronic corpora for tracking large-scale changes such as word order shifts. In other words, the present study aims to test whether the databases used in this research could be used to present feasible arguments on the topic. With numerous factors influencing the distribution of word order (e.g. Davis 2006: 82), some of them can be isolated and compared in the textual data contained in YCOE and PPCME2. Apart from neatly provided dialect information and dates of compositions (or manuscripts), which will enable a plotting of the change in space and time, the corpus material offers the possibility to investigate of word order configurations with respect to particular clause types (matrix versus subordinate), taking into account individual components of word order sequences. The distribution of (S)VO will be analysed to account for different text categories (genres) and whether a given MS is a translation from a foreign original. In order to expedite efficient data handling and analysis, additional statistical tools and techniques have been employed. Among these, the use of the coefficient of variation (CV) provides a complementary angle to the input offered by the standard frequency-based approach.

5.2 METHODS

5.2.1 Comparative approach, theories and models selected

The current study investigates the path of the emergence and stabilisation of SVO order in early English, ensuing (among others reasons) as a consequence of the loss of inflections, which proceeded at a much faster pace in the dialect regions where the impact of medieval Anglo-Scandinavian language contacts was felt most. Morphological simplification would subsequently induce a greater reliance on word order, where caseless noun phrases (i.e. subjects and objects) could still have syntactic function clearly signalled by their specific, fixed position (e.g. McMahon 1994: 122). Should the structural change in question be viewed as a consequence of this particular contact, the diachronic-historical approach has to be considered: the system needs to be, using the words of Manfred Görlach, "compared before and after the transfer" (Görlach 1997: 139).

In accordance with approach, the present research has examined the extent of the preference for SVO in English at two distinct stages of its development: the 'before' stage, viz. Old English, when the presence of inflections would allow for a relative syntactic 'freedom' and the 'after' stage, i.e. Middle English, at which point the lin-

guistic system, with a heavily reduced morphology, would opt for a uniform and, at the same time, more rigid syntactic layout. Further, the exploration of two distinct stages, two distinct synchronic grammars (Görlach 1997: 21) enables a study of the small-scale regional variation which could be a manifestation of the wholesale change (cf. McMahon 1994: 249). This is where the social context comes in, where linguistic data has to be set against the reality of the contact setting at issue and the social reality which subsequently operated. The comparison of synchronic grammars allows a thorough investigation of the distribution of the feature not only in the contact exposed areas but also in all other regions. The multiple perspective option, i.e. checking the distribution of the feature from the many angles adopted in this study, additionally ensures that the effects tracked in the data do not portray an isolated shift but reveal a significant change. In all, an overlap between diachrony and synchrony has to be accounted for. As the corpus analysis in the following chapter will show, the two dimensions can be explored through the same prism.

Comparison between the two language stages is carried out here relying on the model which combines aspects of dialectology and diffusion, viz. Johannes Schmidt's wave theory (*Wellentheorie*; see Schmidt 1872). The framework, thus, incorporates matters of variation and gradual change. The 'wave' suggests that innovations come about in one variety and, subsequently, cross over dialect or language boundaries, affecting others whether related or not. The pace of the spread may be conditioned by socio-political factors (e.g. McMahon 1994: 229-230). This model has been selected not only for the manner in which it portrays the path of linguistic change but also because its aspects can be actually explored in the historical data used in the current research, viz. time (MS dating) and space (dialect information). Adopted, mostly, to account for sound changes, which were at the root of the inflectional erosion in early English, the theory is set to test whether the process of stabilisation of SVO order could re-enact the route paved by the loss of inflections. Particular factors known to affect the distribution of word order sequences (e.g. Görlach 1997: 25 on translation, Swales 1990: 41 on genres), which are marked in the texts of the corpora, also help to bring up specific aspects of the theory chosen for the current research. Methodologically, therefore, the present study constitutes a corpus-based, diachronic as well as synchronic, comparative study of the emergence of the (S)VO word order in early English, brought about potentially as a result of language contact, which promoted inflectional erosion occurring prior to (as well as during) the syntactic reanalysis (cf. Tognini-Bonelli 2001: 65). The quantitative and theoretical formulas are used in tandem to track grammatical change (cf. Fischer *et al.* 2000: 29). It needs to be borne in mind, however, that the adoption of the wave model does not endeavour to fully explain the word order change in English as a phenomenon occurring exclusively due to morphological simplification. Nor does it presuppose that the two processes would share the change mechanisms.

5.2.2 Factors affecting the distribution of word order

In order to address the issues raised in the research aims and to find parallels between linguistic theory and real language data, several comparison perspectives have been adopted (table 5.2.2a below). Some of these (e.g. diachronic and synchronic analyses)

have been motivated directly by the research questions. Others, such as comparing SVO word order patterns including / excluding empty categories or examining the distribution of the feature within different genres, have been included predominantly because of the nature of the two databases used in the current study. The quantitative part of this research, presented in chapter 6, will combine various comparison perspectives to achieve a fuller picture of the issues raised in the research questions and to find answers to them.

Table 5.2.2a: Analysis of distribution of SVO: Comparison perspectives

LEVEL OF COMPARISON	**ASPECT COMPARED**
Diachronic syntax	Old English vs. Middle English
'Diachrony' in synchronic grammars	Subperiods in linguistic epochs – earlier vs. later material (mainly in ME)
Dialectal variation	Dialect sectors of a given corpus (OE and ME)
Parsing scheme	Sequences including/excluding empty categories
Formal grammar – structural representation	(Pro)nominal NPs vs. the non-pronominal option
Formal grammar – syntactic levels	IP- (sentence) vs. IP-MATs vs. IP-SUBs
Syntax vs. stylistic variation	'Genre bias'
Syntax vs. translation	Samples excluding or including translations from foreign originals

Firstly, the preference for SVO will be examined diachronically, comparing the Old and Middle English periods. The expected result is a substantial frequency growth for the investigated feature in Middle English. At this point, the older syntactic conditions ceased to prevail and the contrast in the distribution volume between the two linguistic epochs should be visible especially in the subordinate (IP-SUB) clause environment. The syntactic change would involve quite a dramatic rearrangement of the elements at this level, from the verb final (SOV) to the prevailing SVO. By contrast, a word order associated with the matrix (IP-MAT) clauses would already invite the SVO layout during the Old English period, although rendered essentially, though not exclusively by means of the V2 rule, which no longer operated a few centuries later (e.g. Traugott 1992: 274-275). The possibility of observing the distribution of the word order pattern for particular clause types provides a substantial advantage to the current research. The matrix versus subordinate clause perspective is crucial to the analysis precisely on account of these two distinct scenarios of arriving at SVO.

Next, the emerging SVO sequence will be examined with respect to its constituent NPs – whether pronominal or nominal. There was a particular positional preference for the former to precede the verb and the later to appear after it at the earlier stages of development of the English language, as mentioned in sections 4.1 and 4.3.1.1. The preference eventually withered with the emergence of new syntactic conditions. Furthermore, the character of internal constituents is of essence also when considering the parsing principle adopted for the two corpora. There are many schools promoting distinct ways with which to represent the content of syntactic trees (e.g. McEnery and Wilson 1996: 44), with particular constituent types present or absent from the parsing scheme. The databases used for the current research have been

annotated in a manner which enables searching for both overt and covert elements. The latter are termed 'empty categories' (ECs) by the corpus creators. The overt and covert elements are treated in the databases in the same way except for the fact that the latter will not contain any lexical material. ECs are included in queries by default. In order to exclude them from searches, a special restricting comment has to be made in a query line. This will collectively take away items such as subjects elided under conjunction (*con*), empty expletive subjects (*exp*), "small pro" subjects[92] (*pro*) as well as traces (*T*). The current study does not attempt to address the impact that each of these categories might have on the analysis. Instead, the generic nature of the restricting comment has been used to mainly differentiate between the output data which incorporates the empty material or ignores it for the purpose of pointing to any significant differences in the frequency count. Since one of the goals of the present study is to explore the potential of the corpora in word order studies, it seems pertinent to show the two distinct distributions of the investigated feature, which in turn might create different interpretations of the results. Although no critical assumptions concerning inclusion or exclusion of ECs will be made due to their individual effect on linguistic structures, recognising these distinct distributions could create a new topic for discussion in future studies (see section 7.2 for details).

The dialectal division of the samples in the material provides another definite comparison angle. By observing the preference of the feature within different linguistic regions, not only was it possible to tell which of these favoured SVO diachronically but also synchronically, revealing aspects of variation. The dialectal divisions correlate with the ethnic ones, particularly when English vs. Scandinavian components are examined. Various studies show that the latter made greater use of a surface VO (e.g. Fischer *et al.* 2000: 175), at which point, the possibility of impact of the old Scandinavian syntax on early English structural conditions without the intermediary morphological change needs to be considered as well. Moreover, for the synchronic perspective, additional (statistical) tools are available. It is necessary to employ an output-filtering device particularly for the Middle English period, when the growing preference for SVO may be hidden behind or be blurred by many interfering factors such as periodisation, uneven distribution of distinct genres of the texts or the presence of translations from foreign originals. All these factors contribute to a distinctive (morpho-)syntactic character within the individual texts. They are also likely to influence one another (e.g. Tognini-Bonelli 2001: 132).

In the case of genres, we are dealing with a specific kind of code extant in a group of related texts. The code, in turn, will determine the choice of particular linguistic structures, including word order patterns (cf. Swales 1990: 41, McEnery and Wilson 1996: 101-103; Lee 2001: 39, Tognini-Bonelli 2001: 59). If the selection of genres in a given corpus were evenly represented in samples, the structural interference reinforced by a particular text type, although visible, would not disturb the overall pattern of distribution of the feature in question, where the only aspect truly differentiating

[92] The corpora annotators refer here, however not exclusively, to true pro-drop. The primary reason behind the introduction of the label is to indicate that the (empty) subject of a clause is not coreferential with that of the previous main clause.

samples is a distinct dialect. However, the data available for the present study allowed creation of samples which contain quite a variety of text types, ranging from more (as well as less) formal religious documents to purely secular pieces, with a particular genre either overtly represented or completely absent from the individual dialect sets. In order to overcome (or rather alleviate) this 'genre bias' various, "more expedient" (Biber *et al.* 1998: 253) sampling techniques were employed to establish the possible impact of text types[93] on the layout of the emerging SVO order (cf. Biber *et al.* 1998: 252). These techniques encompassed strategies such as controlled (multiple) sampling, comparison of text doublets[94], stratified sampling see, e.g. McEnery *et al.* (2006: 20) which involved reconfiguration of the output by larger genre groups (i.e. prototypical text categories - section 5.3.4.3) and, finally, a classic comparison of sample versus 'population'[95] (McEnery and Wilson 1996: 64).

Using a wide variety of sampling techniques was necessary in order to estimate the impact of the genre factor, existing alongside other influencing aspects, such as the presence of translations and the (sub)periodisation of texts. Indeed, both corpora include quite a number of translated texts, providing yet another interfering element when investigating the pattern of the distribution of the novel SVO order (cf. McEnery and Wilson 1996: 58, Tognini-Bonelli 2001: 132-156). In the Old English database (YCOE) there is a group of texts that were translated from Latin. As for the Middle English period (PPCME2), a portion of texts translated from (Old) French has been added to the pool (cf. Hogg and Denison 2006b: 34). In both cases, there is a chance of intrusion from a foreign syntactic layout. Latin was predominantly SOV (cf. Fischer *et al.* 2000, 129). Otherwise, the word order would be considerably free with the two unmarked SOV and SVO orders both present (Maiden *et al.* 2011: 401-2). By contrast, Old French would still lean on the V2 principle, with a frequent verb final preference at IP-SUB level (cf. Posner 1997: 357). With new structural conditions, visible especially during the second half of the Middle English period (cf. Fischer *et al.* 2000, 162) the syntactic input introduced by translations would surely distort the layout of the investigated feature. Consequently, a distinction was made between the samples including or excluding most of the translated material (section 5.3.4.2).

Finally, there is the (sub)periodisation of manuscripts, which would affect the distribution of SVO, particularly during the Middle English period. The earlier part of the Middle English material contains copies of Old English works (Fischer *et al.* 2000: 31, Hogg and Denison 2006b: 34), with old and new syntactic tendencies present side by side. To make the investigation more challenging, the documentation from earlier ME is notably meagre (Laing 2000 in McWhorter 2009: 103-5). By contrast, the second part of the epoch is well represented and there is a visible preference for SVO across the dialectal spectrum (Fischer *et al.* 2000: 162 and 175). The distinction between the earlier and later sources, therefore, needs to be taken into account (section 5.3.4.1).

---

[93] The terms *genre* and *text type* have been used interchangeably in the present study, following Stubbs (1996: 11), "in common with most other linguists" (Lee 2001: 41).
[94] I.e. instances of dialect translation.
[95] I.e. the collection of all the texts available.

5.3 DATA

5.3.1 Corpora used, representativeness of material

This study is based on material drawn from two databases: *The York-Toronto-Helsinki Parsed Corpus of Old English Prose* (YCOE) and *The Penn-Helsinki Parsed Corpus of Middle English, 2nd edition* (PPCME2). Both corpora are a part of a larger project at the University of Pennsylvania and the University of York to produce syntactically annotated corpora for all stages of the history of English.[96]

YCOE is a 1.5 million-word corpus of Old English prose texts. As a sister corpus to PPCME2, YCOE uses the same form of annotation. In this respect, both databases can constitute a single unified corpus for diachronic analysis. Still, there are some notable differences between the two. The major distinction is due to the inflected nature of Old English, which required a number of changes to the tagging. The second corpus, PPCME2, contains 1.3 million words, which altogether gives almost a 3 million words database of early English. The Middle English section consists of both poetry and prose. With some additions and deletions, the two corpora rely predominantly on the material included in the *Helsinki Corpus* (HC)[97].

As for the issue of representativeness, the goal of the queries on the two corpora was to first account for variation. Therefore, the material that the samples were generated on needed to represent the actual English language at its early stages with all its diversity. Both YCOE and PPCME2 have the parameters necessary to account for such diversity, including texts from a wide array of dialects, genres and subperiods (cf. Biber *et al.* 1998: 248-50). What is more, the two databases comply with the rules of representativeness also when it comes to the appropriate size of the samples. The current study examines the layout of a word order pattern that is characterised by high frequencies. Thus, the samples need not be extremely large to ensure a reliable quantitative study. The high frequency features are stable in their distribution (McEnery *et al.* 2006: 20, Biber *et al.* 1998: 250, McEnery and Wilson 1996: 64). There is a possibility of slight underrepresentation as far as the number of texts to account for variation across sampling units (dialects) is concerned. Still, the character and stability of the content of these units has been tested along with the extraction of the data from the largest data chunks, when other important factors were examined, such as the interference from foreign translations or the variation across different genres. Finally, much as any annotated corpora, both YCOE and PPCME2 were designed solely to facilitate automatic searching for syntactic constructions. They do not provide an exact, correct linguistic analysis of each sentence.

5.3.2 Sampling frame for YCOE and PPCME2

The sampling frame, i.e. the entire population of texts from which the samples were created, is based on the prose part of the early English corpus (cf. McEnery and

[96] The project was directed by Anthony Kroch, Professor of Linguistics, and the research associate in charge of corpus annotation was Dr. Beatrice Santorini.
[97] The *Helsinki Corpus of English Texts: Diachronic and Dialectal*, a computerised collection of extracts of continuous text, which contains a diachronic part covering the period from c. 750 to c. 1700.

Wilson 1996: 64, Biber *et al.* 1998: 248). While for the Old English database (YCOE) the frame was predefined by its creators, the Middle English corpus (PPCME2) had to be trimmed, excluding items containing verse (e.g. *the Ormulum*). The preference for prose rather than poetry has been determined by the general agreement that data from the former "reflect the language of the speakers most closely" (Fischer *et al.* 2000: 31, Fischer 1992: 209, cf. Stockwell and Minkova 1991: 374). The current study aims at tracking the change in syntax spreading through interdialectal contact, which has the speakers acting as both instigators and propagators of changes. The prose corpus should therefore, at least to some extent, compensate for "the characteristics of spoken language in the historical period" (Biber *et al.* 1998: 252, cf. Davis 2006: 89, Horobin and Smith 2002: 33). Moreover, in Old and Middle English, the written language, indeed, appears to have been closer to how people communicated with each other (e.g. Knowles 1997: 71). It often made heavier use of parataxis than of hypotaxis, contrary to present-day English, where the situation is reversed (e.g. Fischer 1992: 287-8). In addition, in the case of verse, there might be numerous stylistic peculiarities to overcome, many transformations not being actually governed by grammatical context. They could take the research focus too far away from exploring sociolinguistic variation in the distribution of the SVO order, which is the theme of this study (cf. Williams 1975: 232, Mitchell and Robinson 1983: 62). As regards the relation between style and syntactic options at particular clause levels, it has been suggested that prose also proves to be more consistent in exhibiting given word orders for main and subordinate clauses rather than poetry (e.g. Koopman 1998: 141). Finally, it is a well-known fact that poetry leans more on OV structures, which constitutes an older state of events with respect to the syntactic layout of early English (Fischer *et al.* 2000: 174).

5.3.3 Dialect sets as sampling units

Clearly provided dialect information constituted a subsequent and, by far, crucial delimiting aspect when choosing texts for analysis. This research aims to demonstrate that the early Anglo-Scandinavian language contact was the major factor behind the introduction and promotion of morphosyntactic rearrangements. Thus, the input on the distribution of the SVO order provided by the two corpora would show whether the link between the external pressures and the internal rearrangements could be valid. Indeed, dialects constitute linguistic maps that directly correspond to the sociogeographic space. Particular dialect sets in both databases, in this respect, constitute sampling units, which establish the boundary of the 'population' of the prose texts selected for this study (cf. McEnery *et al.* 2006: 19).

The information on dialect in YCOE came from the *Helsinki Corpus* and was given only for the texts that had been included in that corpus. There were some cases in YCOE where the information would not be available; in others, manuscripts would consist of text sections from different dialects (represented by the symbol "/", e.g. West Saxon/Anglian). The former type of texts were not used at all; the latter were incorporated to compensate for the lack of texts in some sampling units (see below). The labels for dialect marking used in YCOE have been listed in the table 5.3.3a:

Table 5.3.3a: Dialect codes for YCOE

DIALECT NAME	ABBREVIATION
Anglian (Northumbrian or Mercian)	A(NM)
Anglian Mercian	AM
Kentish	K
West Saxon	WS
Unknown	X[98]

Since there are no instances of text marked as Northumbrian in YCOE, a division was made within the Anglian dialect section into clearly indicating Mercian and the other, non-specified Anglian (NM), containing texts of mixed dialect[99] one of which was Anglian or chunks representing Anglian cut out from larger files, such as the <codocu.> series which contains charters and wills.[100] Three sets of charters and wills have been modified in this manner:

<codocu1.o1> – only the Kentish part used
<codocu2.o2> – Anglian Mercian and Kentish fragments used
<codocu3.o3> – Anglian and Kentish sections included.[101]

If the remaining part of the chunk was not included in the output tables, it was on account of too small a size, viz. less than 25 clauses, which constituted an optimal minimum for testing. Otherwise, it was a mixed dialect case of West Saxon (over-represented in YCOE) and some unidentified dialect, the inclusion of which would blur dialectal boundaries. Finally, the two Vercelli Homilies manuscripts, the earlier and lengthy, A (with the file name <covercer>) and the later and shorter manuscript E (<covercerE>), originally included in the YCOE without dialect information, have been assigned[102] to the Kentish sector, following Scragg (1992).[103] The complete list (and description, including file names) of the texts which made the samples can be found in Appendix 1A.

[98] For texts including a mix of dialects one of which has not been specified, e.g. West Saxon / X.
[99] Texts of mixed dialect spoke in favour of the one other than West Saxon. The OE corpus has a notable underlying West Saxon tint; therefore, any other dialect marking is significant for the study.
[100] The nature of parsing permits the dissection of the original parsed file as long as the parsing principle is respected and the relevant portion is extracted along with opening and closing codes.
[101] For file name changes upon extraction of text portions in different dialects, see appendix 1A.
[102] I would like to thank prof. Matti Kilpiö of the University of Helsinki for his expert assistance on the dialect assignment.
[103] The main dialect of Vercelli Homilies is, of course, West Saxon. However, there are many substantial 'intrusions' from other dialects as well. According to Scragg, the character of MS. A points to the scribe writing in Kent (although the Kentish forms are not used consistently) (Scragg 1992: f.2 lxxi). The origin of version E, too, should be sought in Canterbury (Scragg 1992: lxxviii). The Kentish origin of the 'Vercelli Book' has been confirmed again by Scragg in his latest work (Scragg in Blanton and Scheck 2008: 369-380). When it comes to the motivation for inclusion of the two versions of the 'same' manuscript in one dialect sector, apart from the overt differences in content (size) and dating, manuscripts A and E were very likely to be drawn from distinct exemplars. Scragg remarks that the scribes were working on the texts of the two pieces independently, with access to a similar set of materials (a library placed firmly in the southeast) (Scragg 1992: lxxvi).

Similar to YCOE, the Middle English database (PPCME2) relies on the *Helsinki Corpus* dialect classification. Additional information has been obtained from LALME (the label 'L' attached to the dialect codes) as well as sources other than LALME (the label 'O' attached to the dialect codes):

Table 5.3.3b: Dialect codes for PPCME2

DIALECT NAME	ABBREVIATION
East Midland	EML, EMO
West Midland	WML, WMO
Northern	NL, NO
Southern	SL, SO
Kentish	KL, KO
Unknown	X

While the portion of Old English dialectal material was overpowered particularly by one regional vernacular, West Saxon, the Middle English database is neatly divided between distinct dialects. However, the content of PPCME2 is not free from problems with respect to dialect 'purity' (Wakelin 1988: 86). One of the impediments is related to the presence of copies and the potential difficulty of identifying the language material as one belonging to the original or as a replica (whether reproduced with or without revisions). Some of the texts included in PPCME2 are doublets available in different dialects. Wakelin (1988) refers to such cases as dialectal 'translations' (1988: 86). The pair of *the Mirror of St. Edmund* (Thornton MS.) <cmedthor**.m34> and** *the Mirror of St. Edmund* (Vernon MS.) <cmedvern.m3> provides a good example. It is a religious treatise written in the Northern and West Midland dialects respectively. Interestingly, the presence of doublets proved significant for the current research. By comparing frequency values for the investigated feature in both versions, it was possible to directly estimate the volume of dialectal influence.

As far as the overlap of the dialect sectors between the two periods is concerned, the table below presents the manner in which the Middle English dialects appropriately correspond[104] to their Old English predecessors:

Table 5.3.3c: Dialect correspondence OE to ME

OLD ENGLISH	MIDDLE ENGLISH
Anglian NM (to compensate for the non-existent Northumbrian)	North
Anglian Mercian	East and West Midlands
Kentish	Kent
West Saxon	South

[104] Following Milroy (1992b: 172, also Wales 2006: 34).

5.3.4 Sample size, dating and other modifications

5.3.4.1 Sizes of samples and (sub)periodisation

For the YCOE corpus, all texts with known dialect information (approximately half of the existing files in the database) constitute the largest sampling units, with dialect sets of unequal size and time frame unevenly spread within them. They cover the span of 456 years, from 743 until 1199. The units include altogether 61 files[105] with the total number of nearly 170000 clauses (IPs) to analyse.

Table 5.3.4.1a: Sampling units containing the largest number of texts in YCOE

DIALECT	NUMBER OF FILES	NUMBER OF TEXTS	TIME SPAN	TOTAL IPs
Anglian (NM)	9	29	940-1099	27,667
Anglian Mercian	5	11	743-1150	16,017
Kentish	6	31	805-1160	7,051
West Saxon	41	196	840-1199	118,075

In order to check whether there are substantial differences in the output values when the time frame is more evenly represented per dialect set, the original, largest units have been clipped to contain 49 files, with the time span beginning from 940[107] and ending with 1199. They make almost 150,000 clauses to search through. These modified units cover the span of 259 years:

Table 5.3.4.1b: Sampling units with modified time span in YCOE

DIALECT	NUMBER OF FILES	NUMBER OF TEXTS	TIME SPAN	TOTAL IPS
Anglian (NM)	9	29	940-1099	27,667
Anglian Mercian	3	8	990-1150	15,940
Kentish	4	25	940-1160	6,825
West Saxon	33	162	940-1199	91,984

For PPCME2, with more uniform time frames, a selection of texts has been made to account for every subperiod within the dialect set. The dating method used in the database is taken from the *Helsinki Corpus*, where a division was made between four distinct (sub-)periods, each covering approximately one hundred years. The appropriate time representation is the key issue especially in the Middle English era, when the shift in the surface word order was clearly visible (cf. Biber *et al.* 1998: 251).

The largest data units contain, for the most part, six files per dialect set, except South which includes 5 files and Kent, where only 3 instances are available in this

[105] Several files in the databases contain many distinct texts.
[106] Dating indicated by a time span has been translated from Latin sigla: Ker's Catalogue of Manuscripts Containing Anglo-Saxon (1956). For charters and wills, instead of a manuscript issue and its date, the authors give the Sawyer (catalogue) number by means of which a manuscript date can be obtained (http://www.esawyer.org.uk/searchfiles/manuscriptsearch.html).
[107] This starting date coincides with the HC divisions (Pintzuk and Taylor 2006: 253)

dialect. These units constitute standard samples used for the majority of comparisons.

Table 5.3.4.1c: Sampling units containing the largest number of texts in PPCME2

DIALECT	NUMBER OF FILES	NUMBER OF TEXTS[108]	SUBPERIODS AVAILABLE	TOTAL IPs
North	6	6	1250-1350 x 2 1350-1420 x 3 1420-1500	8,835
East Midlands	6	6	1150-1250 x 2 1250-1350 1350-1420 x 2 1420-1500	12,758
West Midlands	7	6	1150-1250 x 2 1250-1350 1350-1420 x 2 1420-1500	29,304
Kent	3	3	1150-1250 1250-1350 x 2	6,855
South	5	5	1350-1420 x 4 1420-1500	13,637

The dating issue, however, could not be exhaustively handled only by assuring the presence of the four subperiods in the sampling units. As already mentioned in section 5.2.2, the PPCME2 texts needed to be clearly divided into what was regarded as the earlier and later material. The former on account of the presence of Old English copies and the general scarcity of the documentation from that time; the latter, as containing more texts and with better symmetry between the date of composition and that of the manuscript. The dividing line between the two was drawn at the point of 1350. Furthermore, in those cases where the date of composition was strikingly different from the date of the manuscript (i.e. they belonged to two distinct subperiods)[109], priority was put on the date of the original. Although the manuscript represented the work of a copyist living during the second half of the Middle English epoch, some of the features, some morphosyntactic conditions, from the older original would still linger in the replica (cf. Wakelin 1988: 86).

5.3.4.2 Translations

Both corpora contain a substantial portion of translations from Latin and (Old) French. Therefore, the original, largest units have been reduced to texts least likely to be affected by foreign syntax. It needs to be borne in mind, however, that a sharp division between translated and non-translated material for the data at hand could not be achieved. A small number of texts with non-uniform status[110] with respect to translation were assigned to the non-translated sectors to ensure an equal size of

[108] There is one instance of text included in the West Midland sector dissected into two halves, one being the continuation of the other.

[109] Only three PPCME2 texts chosen for analysis had dates of original and manuscript designating two distinct subperiods. For details, see Appendix 1B.

[110] I.e. texts where translation status is unknown, spurious, often involving multiple 'hands'.

sampling units. Furthermore, some files containing different texts were marked as both translated and not translated, in which case the status of a larger text chunk determined the entire material as either. As untoward as these strategies may seem, we are still dealing with more translated texts versus less translated texts rather than with a clean-cut selection of translations versus native material.

For YCOE, the exclusion of translations substantially abridged the content of the samples, leaving only a little over half of all the clauses available for analysis:

Table 5.3.4.2a: Sampling units in YCOE after the exclusion of translated material

DIALECT	NUMBER OF FILES	NUMBER OF TEXTS	TIME SPAN	TOTAL IPS
Anglian (NM)	2	20	975-1016	5,823
Anglian Mercian	3	4	743-1010	3,451
Kentish	6	25	941-1160	7,051
West Saxon	27	172	840-1199	65,051

Similar to the Old English corpus, the exclusion of texts least affected by translation in PPCME2 reduced the amount of text available in units by roughly fifty per cent (table 5.3.4.2b). The Kentish sector was left unrepresented, as all the texts written in this dialect are, in fact, translations from foreign originals.

Table 5.3.4.2b: Sampling units in PPCME2 after the exclusion of translated material

DIALECT	NUMBER OF FILES	NUMBER OF TEXTS	SUBPERIODS AVAILABLE	TOTAL IPS
North	3	3	1250-1350 1350-1420 1420-1500	4,763
East Midlands	3	3	1150-1250 1350-1420 1420-1500	3,161
West Midlands	4	3	1150-1250 1350-1420 1420-1500	25,303
Kent	n/a	n/a	n/a	n/a
South	3	3	1350-1420 x 2 1420-1500	7,679

5.3.4.3 Sampling strategies for genres

The problem of estimating to what extent a particular text type would affect the distribution of the investigated feature largely concerned the Middle English corpus. Here, the intersection of other influencing factors such as subperiodisation and the presence of translations could be more clearly distinguished and, at the same time, more easily controlled. It was vital to ensure that the interference from these other aspects, along with the 'genre bias', would be taken into account side by side. Indeed, the nature of PPCME2 made it possible to adopt a variety of sampling techniques which would help assess the representatives of the material forming a given sampling unit.

PPCME2, as already mentioned in section 5.3.1, for the most part, contains texts previously found in the *Helsinki Corpus*. The information on the genres is provided by the PPCME2 compilers, often amended with the additional information from the *Helsinki Corpus*:

Table 5.3.4.3a: PPCME2 and genres

TEXT TYPES IN PPCME2
law
document
handbook: astronomy
handbook medicine
handbook other
science medicine
philosophy
homily
sermon
rule
religious treatise
preface/epilogue
proceeding deposition
history
travelogue
biography of saint's life
fiction
romance

Table 5.3.4.3a shows a large array of text types available in the Middle English database. Unfortunately, when building sampling units, a specific genre was either over- or underrepresented, or completely absent from a given dialect sector. Although the current study investigates a high frequency feature, the limited availability and the uneven spread of particular text types within the sampling units could still affect the stability of the output. In order to alleviate the lack of symmetry between the dialect sectors and to enable the estimation of the volume of genre impact on the distribution of SVO between the sampling units, individual genres have been merged into larger entities. They are referred to as "diachronic text prototypes" or prototypical text categories (PTC), following the classification proposed by Rissanen *et al.* (1993) and originally designed for the *Helsinki Corpus* material (1993: 13-14):

Table 5.3.4.3b: Prototypical text categories in PPCME2

PROTOTYPICAL TEXT CATEGORY	GENRES/TEXT TYPES
Statutory (STA)	law, document
Secular instruction (IS)	handbook, science (astronomy, medicine), philosophy, educational, treatise
Religious instruction (IR)	religious treatise, homily, rule, preface, sermon
Expository (EX)	science (astronomy, medicine, other), educational treatise
Non-imaginative narration (NN)	history, biography (saint's life, autobiography, other), religious treatise, travelogue, diary
Imaginative narration (NI)	fiction, romance, travelogue, geography

Grouping the genres into PTCs (table 5.3.4.3b) allowed, first and foremost, for a controlled stratified sampling, where individual sampling units would belong to a single, fairly homogenous group of text types, a genre stratum (McEnery *et al.* 2006: 20). In this manner, the output obtained from the stratified samples could subsequently be compared with the results obtained from larger, more varied units. The availability of sources in PPCME2 enabled the arranging of dialect sets into a unified genre block, representing items belonging to the IR category (religious instruction). The sampling units, thus, contain various kinds of homilies, sermons and religious treatises. In addition, having the division into text prototypes at hand, it was possible to perform the stratified sampling on an entire population. A single dialect sector, therefore, has been further cut into various genre groups. In this respect, the distribution of the investigated feature as conditioned by a particular genre type could be observed in a harmonised dialectal environment (population), free from the interference of other dialects. In the Middle English database, the dialect with the largest number of texts, making a large population, was the one representing the East Midlands. Out of the 25 files available in that dialect, four distinct prototypical text categories were created – IS (secular instruction), IR (religious instruction), NN (non-imaginative narration) and NI (imaginative narration). By observing the individual frequencies within these categories, the effect of particular prototypical group onto the distribution of the SVO order could be estimated.

Having a large number of texts from a single dialectal set additionally enabled execution of two other sampling strategies. One, a standard sample-to-population comparison, was used to assess to what extent the content of a chosen sample would be representative of the entire population. The output from the sample taken from the East Midland set would be, thus, set against the results generated from the entire material available in that dialect. It was not only the impact of various genres that was examined but the nature of the population calqued the image of the sample when it comes to other important intervening factors such as subperiodisation and the presence of translations. The second technique, also estimating the representativeness of the sample used for main research comparisons, involved multiple sampling from the same population. In order to estimate the stability of the generated output, the results from particular 'sample variants' would be set against one another. The issue of an uneven distribution of genres would be set aside in this case. However, the

restriction as to the choice of material would be observed, this time, with respect to subperiodisation. Each sample produced came to contain the same amount of distinct files from given subperiods.

5.3.5 Data retrieval: query language, sequencing and handling data

Corpora queries[111] were generated to fulfil the conditions of the research hypothesis where a connection between the case erosion and the growing prevalence of SVO order was checked. The SVO sequence was realised as the NP – VP – NP configuration, to follow the corpora annotation. It contained only the (originally) inflected content of each of the elements in the configuration. The linear order presents the rigid formation which began to operate once the inflections withered and the grammatical functions could no longer be read off the morphological endings. The first NP in the sequence had to be nominative, therefore the subject, and the second would be oblique, the object. It is the formation which has been observed in English until today, as presented in section 3.1.

5.3.5.1 Subjects, Verbs, Objects in YCOE

In the Old English corpus, regardless of the neat parsing principle adopted, the representation of what was to become the subject and object had to be further defined. Unlike in PPCME2 where subjects and objects were recognised by a clear-cut annotation, there were case labels introduced in YCOE both at word (POS tagging) and phrase level (PSD tagging). With subjects, any NP with the nominative tag (NP-NOM) realises the subject in the sequence. In the case of objects, the NP possibilities making the sentence element had to be listed in a special definition file. From the definition file the search programme would, in turn, acquire the information as to which labelled item to include in the selection of sequences. The object was, in this respect, described as any of the following noun phrases:

Table 5.3.5.1a: NPs classified as objects in YCOE

PSD LABEL	ITEM REPRESENTING
NP	Uncased NP
NP-ACC	NP in accusative
NP-GEN	NP in genitive
NP-DAT	NP in dative
NP-RFL (NP-RFL-ACC, NP-RFL-GEN, NP-RFL-DAT)	NP reflexive, cased argument
NP-RSP (NP-RSP-ACC, NP-RSP-GEN, NP-RSP-DAT)	NP resumptive, cased argument

Some NPs in the data are uncased. Fortunately, there was no confusion as to what eventually would make a subject or an object. Contrary to objects, nominative NPs are always identified by case, at least at the phrasal level, because, as in PPCME2, subjects in the Old English corpus have a special status. Every complete finite clause

[111] For the examples of queries used in this research, see appendix 3A and 3B.

has a subject (overt or empty). It is always possible to distinguish the subject argument from all others.

When accounting for a particular verb phrase (VP) type, the entire element needed to be constructed accordingly[112]. Out of a wide array of possibilities, the most effective way to build up a complex VP was to again make use of a definition file, which would comprise all the potential verbal candidates. With the YCOE annotation, a distinction was made between finite and nonfinite verbs:

Table 5.3.5.1b: VP elements in YCOE

VERB TYPE	POS LABELS	ITEM REPRESENTING
Finite verb	MDP, MDD	Modal, present or past tense
	HVP, HVD	The verb HAVE, present or past tense
	BEP, BED	The verb BE, present or past tense
	VBP, VBD	All other verbs, present or past tense
	AXP,AXD	Auxiliary verb, present or past tense (indicative or subjunctive)
Nonfinite verb	VB, VBN, VAG	Lexical verb, infinitive (inflected)
	VBN	Past participle
	VAG	Lexical verb, present participle
	HV	HAVE, infinitive (inflected)
	HVN	HAVE, past participle, verbal or adjectival
	HAG	HAVE, present participle
	BE	BE, infinitive (inflected)
	BEN	BE, past participle

5.3.5.2 Subjects, Verbs and Objects in PPCME2

For the Middle English database, the only element in the sequence which needed to be defined was the verb (phrase). Subjects and objects, as mentioned earlier, had clear –SBJ and –OB function labels attached to any of the NPs. For the VP constituents, similar to the YCOE classification, particular definition files included separate lists of what was to be treated as a finite and a nonfinite verb:

[112] A separate VP node was not used in the two corpora as the surface order of the VP in Old English was still relatively 'free'. In this respect, determining the boundaries of this particular node would be difficult and time-consuming. As a result, the corpus compilers decided to treat the verb (or verbs) and all its arguments (both subjects and complements) as sisters, directly dominated by the sentential node (IP).

Table 5.3.5.2a: VP elements in PPCME2

VERB TYPE	POS LABELS	ITEM REPRESENTING
Finite verb	MD	Modal, present or past tense
	HVP, HVD	The verb HAVE, present or past tense
	DOP, DOD	The verb DO, present or past tense
	BEP, BED	The verb BE, present or past tense
	VBP, VBD	All other verbs, present or past (indicative or subjunctive)
Nonfinite verb	VB	Infinitive, verbs other than BE, DO, HV
	V*N	passive or perfect participle
	HV	HAVE, infinitive
	H*N	HAVE, passive or perfect participle
	DO	DO, infinitive
	D*N	DO, passive or perfect participle
	BE	BE, infinitive (inflected)
	BEN	BE, past participle

There are differences between YCOE and PPCME2 in the annotation of the VP element, with the notable addition of DO in the latter. DO was an all-purpose function word in Middle English, as opposed to Old English where it functioned as a main verb, hence no distinction in the YCOE labels (cf. Williams 1975: 272).

5.3.5.3 (Non-)pronominal option

This study explores the extent of Scandinavian influence on early English and thenceforth. One of the examined aspects involves the consequences of morphological simplification on the subsequent alterations and the growing rigidness of word order. The other encompasses the changing status of personal pronouns, which from "a much more pronounced surface differentiation" in OE (Fischer *et al.* 2000: 71) began to align with the structural behaviour of full NPs during ME. The result of both developments was the arrival of the strict SVO. In both cases, contacts with the early Scandinavian population have been regarded as potentially instrumental. Consequently, both change scenarios had to be addressed and explored.

Having fitted corpus queries to investigate the morphological simplification angle, viz. (S)VO which included all constituents of NPs, consecutive query line modifications had to account specifically for the presence or absence of pronouns in the element strings. In this respect, a separate set of searches has been generated to check the distribution of SVO with NPs which encompassed solely (iDoms [1][113] PRO) or entirely excluded (iDoms [1] non-pronominal_NP[114]) personal pronouns, acting as subjects or objects. However, the default option for SVO strings including both nominal and pronominal elements has been used regularly throughout the data analysis.

[113] Arguments that occur more than once in a query line are given numbers so that CSearch can recognise them as separate entities.
[114] For non-pronominal option, a definition file had to be created (appendix 2B).

5.3.5.4 Element sequencing

When constructing the SVO sequences for the queries, the closest proximity of elements was emphasised. In this respect, the search function 'iPrecedes' was employed. It did not allow for any interruptions in the string and the output produced comprised only the items specified in the query line, S + V + O. However, there was a number of element combinations[115] to be taken into account with respect to the position of the finite and nonfinite verb as well as the negative particle (NEG)[116] when building up the SVO queries both for YCOE and PPCME2. Further, the shift from Old to Middle English syntactic patterns involved, among others, a steady increase in verbal periphrasis along with a substantial change in the frequency of discontinuous verb phrases. In this respect, many of the candidates for SVO sequences, especially those retrieved from the YCOE samples, could potentially include only a part of a discontinuous verb phrase, giving rise to potential problem in calculation. Consequently, in order to alleviate the problem and to control the distribution of the SVO order between the two periods, the incomplete element strings had to be subtracted from the output[117].

In addition, a decision had to be made with respect to adverb phrases (ADVP). They were excluded from the SVO strings. There is a distributional discrepancy between the northern and southern dialects in ME as far as the sentence initial adjuncts are concerned. The South allowed them in sentences whereas the Scandinavian influenced North did not use them (e.g. Mitchener 2005: 11). A series of control searches[118] on SVO patterns with or without ADVPs in the element string, indeed, revealed a marked preference for the investigated feature with ADVPs for the southern dialects (Kent, South) visible at particular clause levels. Finally, as mentioned in 5.3.5.3, a (non-)pronominal variant was added to the sequencing option in the queries for PPCME2. The nominal versus pronominal distinction on noun phrases was made on account of the change in the positional preference for pronominal objects with respect to the verb (e.g. Williams 1975: 232). However, the sequences with non-pronominal subjects were also examined. In general, to ensure the comparability between the two corpora all the available sequencing rules had to apply to both databases at the same time. In other words, what was valid for Old English would still need to be searched for in Middle English, and vice versa.

One of the SVO strings with a complex verb element as retrieved from YCOE is presented below:

---

[115] Potential sequencing options have been confronted with Williams (1975: 230-2 and 270-1) and Fischer *et al.* (2000: 143 and 145).

[116] Although the status of the negative was far from fixed at the two periods, with considerable variation in the placement and number of the particles used, NEG was added to the SVO sequences, as the close connection between the negative and the main verb had never been disturbed (cf. Fischer and van der Wurff 2006: 157).

[117] The problem of the faulty retrieval of fragments of discontinuous verb phrases by CSearch could not be fully resolved. The programme could only pick the latter fragments of VP within the closest vicinity of the searched structure. An example of error estimates for PPCME2 can be found in appendix 5A. Although the number of 'bad' tokens has been recorded, no major deviations from both the average frequency values for dialect sectors as well as from the general distribution pattern has been detected.

[118] Tables comparing sequences with and without ADVPs are presented in appendix 5B.

Table 5.3.5.3a: SVO pattern as extracted from YCOE

Plain text	**His leorningcnihtas woldon gelettan þone Hælend, …**
Modern translation	His disciples would/wanted to stop/hinder/delay the Saviour, ...
Fragment coding	(coaelhom,ÆHom_6:334.1032)
Parsed text	((IP-MAT (NP-NOM (PRO$ His) (N^N leorningcnihtas)) (MDDI woldon) (VB gelettan) (NP-ACC (D^A +tone) (N^A H+alend))

For PPCME2, with a distinct annotation for subjects and objects, an example of the SVO sequence is as follows (with a negative particle placed between the finite and nonfinite verb):

Table 5.3.5.3b: SVO pattern as extracted from PPCME2

Plain text	**… but he mi3t nou3t smyte hit.**
Modern translation	… but he might not destroy it.
Fragment coding	(CMBRUT3,29.858)
Parsed text	((IP-MAT (CONJ but) (NP-SBJ (PRO he)) (MD mi+gt) (NEG nou+gt) (VB smyte) (NP-OB1 (PRO hit)) (. .)) (ID CMBRUT3,29.858))

5.3.5.5 Handling the data

Since the present study deals with word order patterns where particular sentence elements form longer strings, the search results needed to be analysed not against the overall number of words but against that of clauses per given text. The output was available, therefore, at three levels: the total of all the matrix and subordinate clauses found in a particular text and at matrix and subordinate levels separately. The comparison of scores between the sentence and matrix-subordinate levels helped to identify which of the latter two specifically contributed to the final values. Most importantly, however, the matrix-to subordinate clause comparison, as mentioned already in section 5.2.2, allowed separating two distinct environments of syntactic change in early English, the one with (matrix) and the other without (subordinate) the V2 rule.

5.3.5.5.1 A frequency-based model and multiple common bases

As with any standard case of corpus research, this study relies on a frequency-based model, whereby the retrieved number of occurrence of the investigated feature in particular sampling units can be used to construct arguments, to accept or reject hypotheses presented on the topic. In order to make valid comparisons between the sets of data, the raw count in each sampling round was filtered through the statistical

significance tests (section 5.3.5.5.3 below). Yet, for an accurate reading of the data at hand, the number of (S)VO tokens had to be further normalised, to adjust counts from texts of different lengths (cf. Biber *et al.* 1998: 263). The lowest common base adopted for this purpose was 25 clauses (IPs) (see e.g. McEnery *et al.* 2006: 53). It enabled the examination of a large number of texts combined from the two databases, with text specimens available in all dialects at the two distinct periods. Subsequently, larger common bases were adopted, viz. 50, 100 to 400 IPs. Having to deal with a high text length inconsistency within the two corpora, it was necessary to examine whether and how frequency differences between particular dialect sets fluctuate for a given common base. By adopting a given base for data handling, the size of dialect samples changed, undoubtedly influencing the results on the distribution of the investigated feature. Furthermore, the multiple common base approach allowed observing whether the overall pattern of changes discussed remained stable throughout the corpora.

5.3.5.5.2 Coefficient of variation

In order to accurately observe the language change in the making, as presented by frequency values in the overlapping dialect areas between the two periods, the output (i.e. normed frequencies) had additionally to be filtered through a statistical measure which would elicit the relevant facts from the results charts. The measure known as the coefficient of variation (CV) was used for this purpose. This measure is defined as the ratio of the standard deviation to the mean. It is a useful statistic for comparing the degree of variation from one data series to another. More specifically, the coefficient measures how much individual output values deviate from the estimated average of the feature found in a particular series. In the present study, the CV shows to what extent the frequencies calculated for particular texts differ from the average frequency estimated for the whole dialect set. The higher the value of CV, the greater the variation (or irregularity). Conversely, the lower the value of the coefficient, the more uniform representation of the feature in particular dialect sectors. The CV is calculated as follows: the ratio is multiplied by 100 and the deviation is expressed as a percentage of the mean (Frank and Althoen 1994: 58-9). When the measures are changed to per cents of their averages, direct comparisons of the representativeness of particular two averages can be made. By contrast, attempting to compare the validity of the averages by means of direct measures of dispersion would not provide the necessary degree of precision in the interpretation of the data (e.g. Balsley 1964: 91-92).

In order to illustrate how the tool works, Figures 5.3.5.5.2a and 5.3.5.5.2b below present the distribution of normed frequencies[119] for the SVO sequence in the northern and southern Middle English dialects, respectively, at the sentence level (for element strings including empty categories). The bold vertical line stands for the average frequency in a particular dialect set. The length of columns on the right and on the left off the vertical line show higher or lower frequency values found in individual texts. The North (Figure 5.3.5.5.2a) scores 6 per cent for CV, showing a regular distribution of the investigated feature – the columns barely deviate from the vertical line (i.e.

[119] The figures are based on the results from the pilot study.

average), as opposed to the bars of the graph for the South (Figure 5.3.5.5.2b), where the departures from the average are more visible, hence the higher CV of 16 per cent.

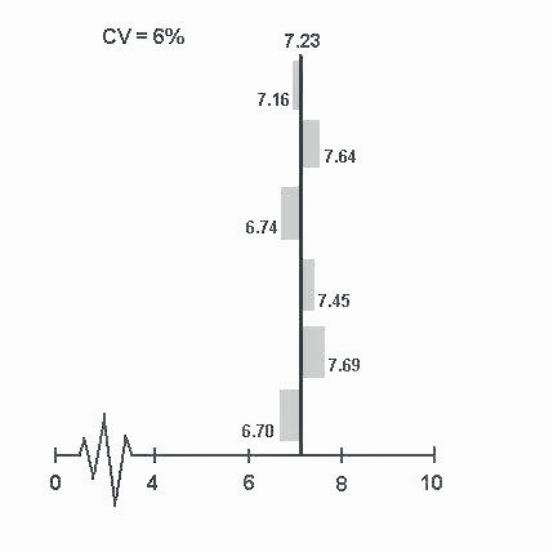

Figure 5.3.5.5.2a: Coefficient of variation for the North - SVO overall with frequencies for individual texts specified

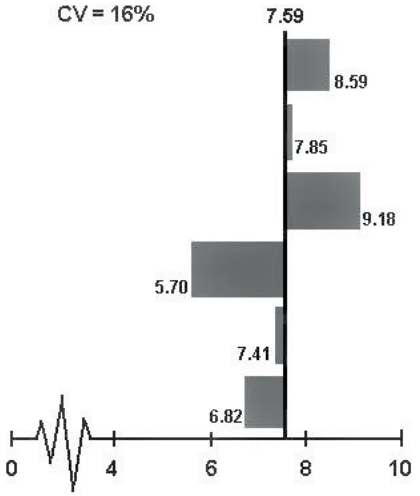

Figure 5.3.5.5.2b: Coefficient of variation for the South - SVO overall with frequencies for individual texts specified

It is important to notice at this point already that relying on the average normed frequencies alone does not suffice to adequately interpret the results. Here the South, as indicated by the average frequency values, scores higher for SVO than the North, but the level of regularity displays an opposite pattern.

The reason for using the coefficient as an additional measuring filter is that it provides much more information about the exact nature and degree of variation in the datasets. The tool proved especially useful when average frequencies in sets did not differ significantly from each other, which occurred frequently with normed frequencies calculated on the adopted common base and especially for the lower common bases. In this respect, CV helped to differentiate seemingly identical outputs (cf. Calmorin 1997: 103). What is more, the coefficient properly translated the figures into facts. As the two databases present an uneven distribution of texts per dialect, per time and per genre, any modification to the length of particular units changed the rate of fluctuation. The coefficient, therefore, controlled the frequency values when discrepancies between the samples were too great to be solely used for the interpretation of linguistic phenomena. Furthermore, apart from shedding light on the nature of the 'real' variation, the coefficient enabled testing the reliability of the output. It showed whether a high condensation of the feature per data set was obtained from a more regularly distributed feature within texts (very low CV), pointing to a high reliability of the results, as opposed to cases where the frequencies of the feature within different dialect sectors fluctuated too much, thus increasing the unreliability of the final output. The values of the CV exceeding 60 per cent equals treading on the borderline between what would be classified as acceptable and unacceptable in terms of valid interpretations (cf. Yamane 1973: 79). In this respect, any section which scored higher than the above threshold value was not used to form arguments in the present research.

5.3.5.5.3 Statistical significance tests

As is customary in frequency-based research, statistical significance tests constitute a mandatory procedure, determining whether the retrieved (sets of) data can be set against one another to enable valid comparisons and interpretations. Since parallel comparisons, involving the investigation of various factors simultaneously could not be made due to the nature of the data at hand[120], statistical significance tests chosen for this study were those which can handle problems with sets of two samples analysed, one set at a time. The software used for this purpose was 'R', a widely used tool for statistical computing and graphics (e.g. Crowley 2007).

For the vast majority of analysed cases a classic chi-squared test was used. It has been designed to check the independence of two variables in a contingency table (Crowley 2007: 354 and 365-368). With some instances, where normed frequencies pointed to a particular sampling unit as the one standing out in the unit table, apparently conditioned by the presence or absence of an aspect compared at a given stage[121], binominal tests comparing two proportions were used (Crowley 2007: 365). Only the comparisons which gave *p*-values below the significance level of 0.05 were considered for interpretation.

---

[120] The impact of genres could be measured more precisely only within one sampling unit, the East Midland dialect. Exploring of the time factor, which involved portions of material from the earlier and later part of ME, was incomplete due to the lack of texts from the first half both for the North and the South.

[121] E.g., the odd contrast in normed frequency values between one dialect sector alone when empty categories are present or absent from the sequencing option.

5.3.5.6 Differences between YCOE and PPCME2

Although both YCOE and PPCME2 are constructed and annotated in roughly the same manner, there are some notable differences between the two databases. The major discrepancies relevant to the present study revolve around the treatment of cases, the internal noun phrase annotation and the tag distinction at the clause level.

In YCOE, case, as emphasised by the compilers, is a major formal category. Sometimes it replaces the function labels of PPCME2 and sometimes it is used along with them. Case is labelled at both word and phrase levels. However, not all inflected words and phrases are inflected for case. The rules for case tagging rely on formal identification and the compilers left certain kinds of ambiguities unresolved, giving rise to an underspecified system of labelling. In this respect, some noun phrases (and other nominal categories) do not carry a case label even though they are, in fact, inflected for case (whether or not the case can be determined from context). The decisions about cases were based on the gender of the noun as listed in Hall's *A Concise Anglo-Saxon Dictionary, Fourth Edition* (1960).

The part-of-speech tag referring to proper nouns in the Old English database is NR whereas in the PPCME2 it is NPR. The distinction makes it possible to avoid the confusions which would occur when using 'NP*' as a search-term, retrieving both regular noun phrases (NPs) and proper nouns (NPRs).

One can identify complete and incomplete clauses more clearly in the YCOE than in the PPCME2. In both databases, a system of equal-sign co-indexing is adopted to point to clauses in which elision has occurred. In this respect, in the PPCME2, a clause label ending in =# (where # designates any number, e.g., IP-SUB=1) is an incomplete clause patterned on a clause with a matching -# index (IP-SUB-1). In the YCOE the same system is used, although for technical reasons the index is always =0, =00 or =000. The clause label of the incomplete clause in YCOE is different as well, being IPX (or CPX) rather than IP (or CP). In the present study, all the incomplete clauses needed to be identified and excluded from the respective IP-MAT and IP-SUB pools. They were not considered in the final analysis of the data.

6 Results

6.1 DISTRIBUTION OF SVO IN THE CORPORA

The current chapter presents the results of the searches on the two corpora, YCOE and PPCME2, for the distribution of the SVO order. The focus has been placed on differences in the dispersal of the feature between the dialects which correspond to the areas of Scandinavian influence (the North and the East Midlands) and the rest of the dialect regions (the West Midlands, Kent and the South). Firstly, the distribution of the feature will be shown and analysed diachronically, using the input from the Old and Middle English data. The two linguistic epochs will be compared, looking at the largest, not-yet-modified samples, later to be confronted with sections which contain more uniform portions of the data, both with respect to subperiod representation and the absence of translations. Further, the preference for SVO will be evaluated within and between particular clause levels, since the changing syntactic conditions at both the earlier (OE) and later (ME) stages of development of English may involve distinct mechanisms on matrix and subordinate clauses, respectively. The subsequent sections, then, deal with the findings concerning the new syntactic situation during the Middle English period, in particular. Apart from the comparison between clausal environments as well as the output evaluation upon the extraction of translated texts, differences in the distribution of SVO between the earlier and later parts of Middle English material will be presented. The distribution of the feature will also be assessed, taking into account the potential influence of distinct genres. After examining the patterns of distribution of SVO in early English, the comparison perspectives just mentioned have also been adopted to check the applicability of a diachronic corpus study to explain the nature of the morphosyntactic changes tackled in the present research. Furthermore, the distribution of SVO will be presented focusing on the nature of its constituent NPs, viz. either pronominal or non-pronominal.

The results will be returned to in the second part of this chapter (section 6.2), in which I will try to consolidate the arguments to address the problems presented in the three main research questions laid out in 5.1. One of the goals of the current study is to assess the correlation between the morphological simplification and the concomitant shift to the rigid SVO. The role of the external pressures, especially Anglo-Scandinavian language contacts, in these two developments is also examined. Yet another aspect to be assessed is the relevance of the widely used theoretical model, the so-called wave theory, to the changes discussed here.

6.1.1 Emergence of a prevalent SVO order: Old through Middle English

The prime reason for inclusion of the OE data in the results has been to establish the baseline which would enable an adequate analysis of the ME data. The focus, therefore, has been placed not so much on particular average (normed) frequency values in the YCOE (and PPCME2) dialect sets. Instead, the attention has been turned towards

identifying the expected overall growth pattern of SVO from OE towards ME.

The SVO order in OE was one of the structuring possibilities, frequently generated with a single finite verb along the lines of the typically Germanic V2 rule, which was the syntactic condition limited predominantly to the matrix clause environment. Indeed, sequences with a single finite verb constituted a vast majority of the retrieved instances from YCOE:

(7) (…) Ond he dyde monig heofonlic wundor, (…) [ID comart3,Mart_5_[Kotzor]:Jy27,A.25.1289][122]
(8) Cristes leorningcnihtas leidon heora reaf uppon þam assan (…) [ID cocath om1,+ACHom_I,_14.1:293.98.2641]
(9) (…) and þa Deniscan ahton wælstowe geweald. [ID cochronC,ChronC_[Rositzke]:833.1.507cochronC]

The number of identified SVO strings was greater in the matrix clause pool than the one recorded at the subordinate clause level.

By contrast, Middle English witnessed an extensive syntactic reshuffle, with the rigid SVO order on a steady increase in both matrix and subordinate clauses. At the same time, the number of other word order possibilities decreased. With the overt case-marking no longer present, it would be increasingly challenging to interpret these (now more marked) word orders accurately. They would, thus, be avoided in the acquisition process, as mentioned in section 3.2. The OE to ME comparison, therefore, has been devised to evaluate the increase of SVO especially with respect to specific clause levels and to estimate the distinct growth rate within particular dialect zones.

When comparing the average normed frequencies in dialectal sets as a whole between the Old and Middle English portions of the material, the distribution of SVO conforms to the expected pattern. As figures 6.1.1a through c show, there is a substantial growth in the subject-verb-object word order, with the volume of increase ranging from 50 to over 80 per cent[123], depending on the dialect and the common base adopted for normalising. As far as particular sampling units are concerned, the largest increase in the preference for the investigated feature takes place in the northernmost dialects. The rate of increase then decreases southwards, one sampling unit by another, until it reaches its lowest point in the southeastern Kent sector. The two dialect sectors which favour the SVO order most are, in fact, those corresponding to the areas where the early Scandinavian presence was attested: the North and the East Midlands. Of these two, the North displays the most substantial increase of the feature. It is, essentially, the very same dialect where inflections are said to have eroded first. However, drawing a direct link between this morphological simplification and the structural rearrangement that followed suit is perhaps, at this point, a little premature. Admittedly, the not so prominent SVO growth pattern for the southern dialects is, at least to some extent, a result of the word order being already present among the early English syntactic tradition, although quite distinct pressures pro-

[122] Instances of sequences extracted from both corpora can be found in Appendix 3A.
[123] An example of the table with the output data used to generate figures in 6.1.1a is included in Appendix 4.

moted this particular structural option during the earliest days of OE. Nevertheless, the substantial growth in the preference for the SVO order within the northernmost dialect sectors should not be underestimated. The rise from 50 per cent up to 80 per cent for SVO is definitely indicative of a very sharp shift towards a new syntactic framework. Intriguingly enough, the very distinct, highest preference for SVO in the North, as will be seen in the subsequent sections, is repeated throughout almost all data handling configurations, with the contrast between this and the rest of the sections clearly visible. Finally, both chi-square as well as proportion tests conducted on various data groups revealed the differences between the sets as either very or extremely statistically significant.

For the overall[124] SVO distribution within the range of common bases from 25 up to 100 clauses (Figure 6.1.1a), there is a familiar growth pattern, going from its northern peak to the southern lowest points. The most orderly distribution of the feature is found with the units normed to the common base of 25 clauses. These units contain the largest amount of various text types taken from many distinct subperiods, especially in the Old English part, ensuring a wide variety within the samples and thus enhancing their representativeness. The SVO growth pattern is maintained here steadily dialect by dialect, from the highest, i.e. the North, to the lowest, Kent.

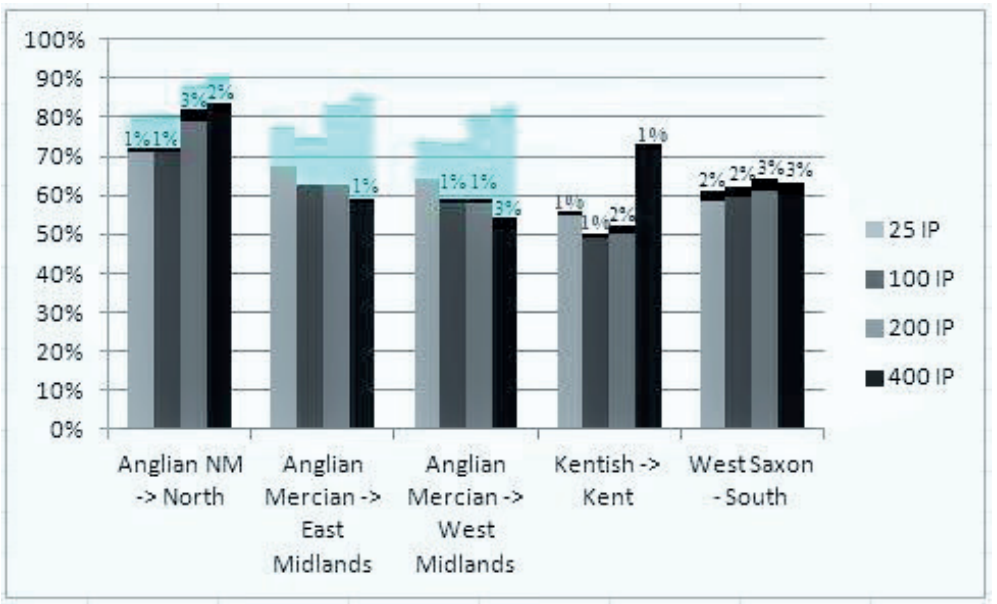

Figure 6.1.1a: Growth of SVO from OE to ME overall for the largest, uncut samples: common bases from 25 to 400 clauses (differences between sequences with and without the empty categories shaded in black, output calculated for Anglian as a uniform sector shaded in light blue)

Similar growth pattern comparisons were made on the basis of higher common bases of 200 up to 400 clauses in order to bring the content of the two databases closer to

[124] I.e. the total of both matrix and subordinate clauses.

each other with respect to the size of the individual texts. The downside of the larger common bases was a clear loss of representativeness in the Old English section. Thus, the YCOE samples lost well over a half of the material, with quite a number of texts (files) being too small in size to be included in all the units except the 'over-represented' West Saxon one. The rising preference for SVO calculated for the common bases of 200 through 400 clauses resembles that of the lower common bases, with the contrast between the sequences including and those excluding the empty categories (black shading in the columns) sharper than the one recorded for the common bases of 25-100 clauses. The biggest growth, again, is seen in the northernmost set, where the distribution volume exceeds 80 per cent. There is a substantial boost in values for the common base of 400 clauses in the Kentish section. This increase is most likely due to the sample modification, which resulted in the loss of representativeness in the unit. A drastic cut was made on the YCOE Kentish sample in order to include those texts which could be normed to such a high common base. In the end, the set came to include only two files, which contained two distinct versions of the *Vercelli Homilies*.

In order to establish to what extent the content of the Anglian NM set, arranged to compensate for the lack of Northumbrian texts in YCOE, could potentially distort the genuine increase rate of SVO (if at all)[125], a separate growth layout was generated to account for possible output inconsistencies and unnatural fluctuations (blue shaded columns in figure 6.1.1a). In this new configuration, the Anglian sector no longer distinguishes between the Mercian and the NM sets but provides the average frequency value for the Anglian dialect as a single, uniform sector, generated from all the texts marked as Anglian, including the mixed dialect cases. The highest preference for SVO is still, expectedly, found in the sampling units corresponding to the areas of Scandinavian influence, i.e. the two northernmost dialects. The differences between particular units are understandably smaller, but the same growth rate pattern is nevertheless maintained.

Let us next turn to the clause-type level and begin with matrix clauses. As figure 6.1.1b shows, the distribution of SVO within these clauses appears, by and large, quite uniform from Old to Middle English. The biggest growth in SVO is found, as could be expected, in the northernmost set, with the differences between the outputs generated for the common bases of 25, 100 and 200 clauses varying considerably (ranging from 70 per cent up to 85 per cent). However, once the texts from the Anglian sector are treated as a single set, the majority of the growth rate columns reach the same level in all the sampling units with the exception of Kent. Due to the sample modification carried out to fit the higher common base, the Kentish dialect again lost almost all texts, causing a considerable rise in the growth pattern when one moves up to the common base of 100 and 200 clauses.

[125] The NM sector contains file clippings from larger multi-dialect text portions. It also includes cases of mixed dialect one of which is Anglian.

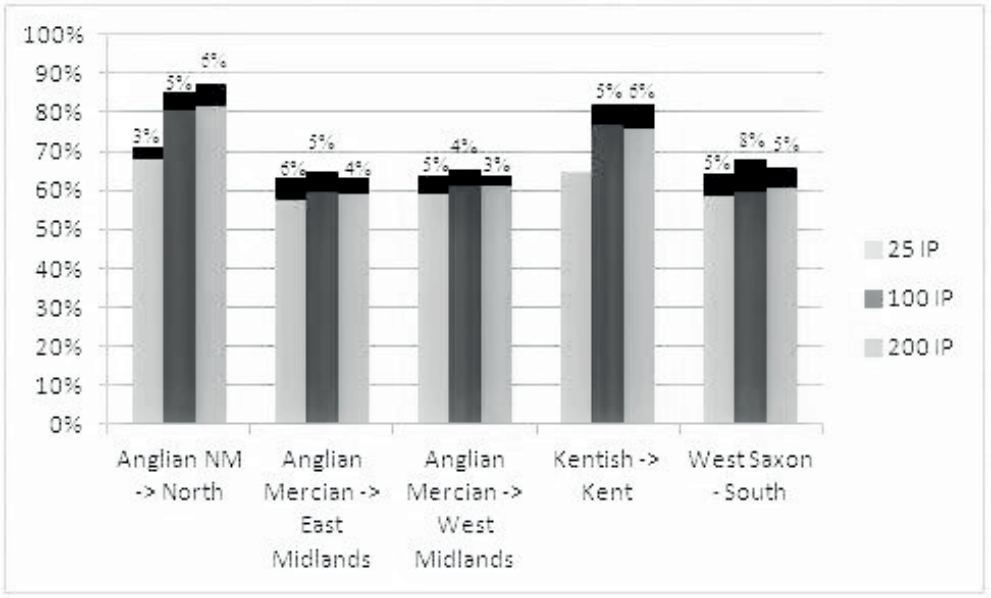

Figure 6.1.1b: Growth of SVO from OE to ME at matrix clause level for the largest, uncut samples: common bases from 25 to 200 clauses

To sum up on the matrix clause level, the more regular growth pattern for SVO in this clause environment could be expected in view of the fact that growing prevalence of SVO merely added tokens to the already existing preference for the SVO order at that clause level.

In subordinate clauses, on the other hand, the SVO growth pattern within subordinate clauses presents a different but, nonetheless, intriguing picture (Figure 6.1.1c). There is a recognisable, highest preference for the investigated feature found, once again, in the northern set. Still, the increase is weakening steadily southwards dialect by dialect. The cascading decline is especially visible for the common base of 25 clauses, where the largest number of different text types is compared. What is more, the outputs for that particular clause range witness the highest increase in frequency not in the northernmost sector but in a somewhat more southerly set, viz. the Anglian Mercian one, that is the East Midlands zone. This set also belongs to the Scandinavian influenced regions. Nevertheless, this time the highest frequencies for SVO are found outside the northernmost dialect areas.

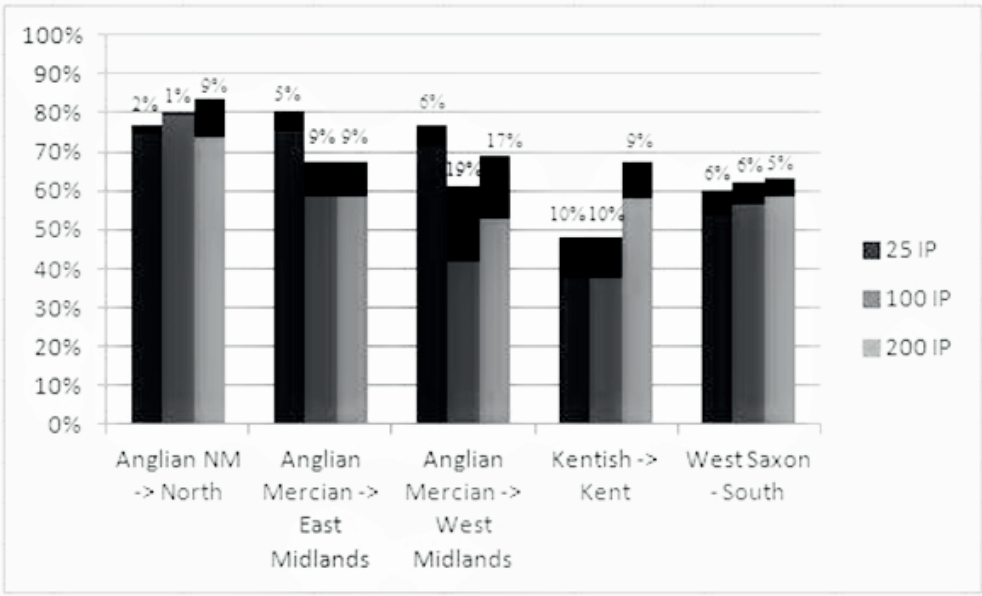

Figure 6.1.1c: Growth of SVO from OE to ME at subordinate clause level for the largest, uncut samples: common bases from 25 to 200 clauses

Interestingly, the output generated from the other common bases follows the recurrent pattern where the highest frequency of SVO occurs in the North, decreasing southwards set by set. Yet, the contrast between the northern set and the rest of the units is, again, a little sharper than the one displayed for the common base of 25 clauses. Predictably, the rise was, at least to some extent, caused by substantial alterations in sampling units (in YCOE) – deletion of texts shorter than 100 or 200 clauses. In all, a more differentiated growth pattern for SVO at the subordinate level implies the development of (or shift to) a new syntactic framework with a greater number of structural transformations (SOV to SVO) when compared with the matrix clause level (frequent SVO often with underlying V2), thus generating a less distinguishable growth rate layout between the dialects.

Next, when observing the increase in SVO with a more uniform distribution of texts in sampling units with respect to the time span investigated[126] (Figure 6.1.1d), the output layout reveals no major divergence from the overall trend.

[126] Time span norming concerned mostly the Old English sectors, where some of earliest texts were removed from some of the sampling units. See section 5.3.4 for details.

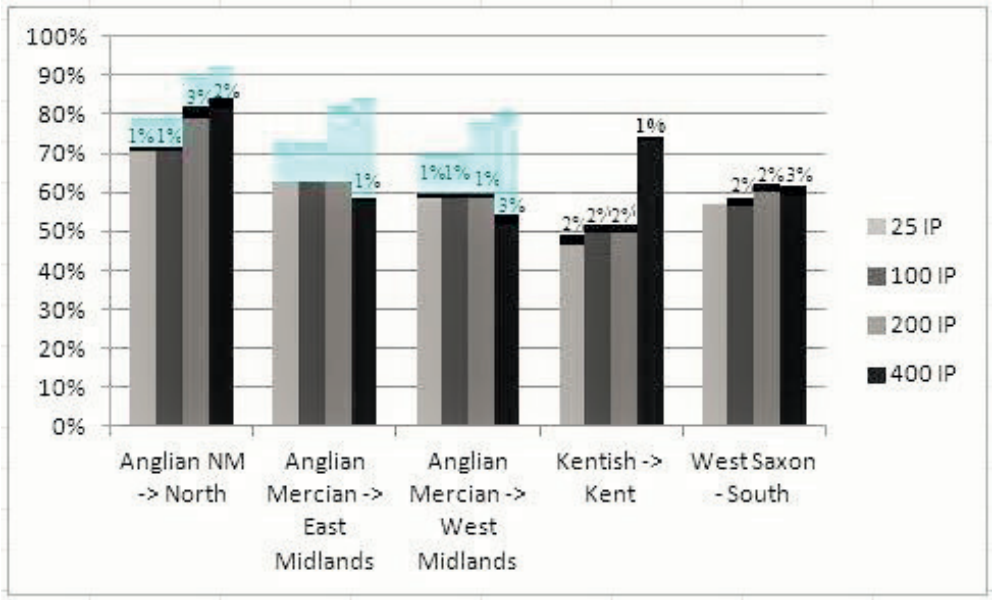

Figure 6.1.1d: Growth of SVO from OE to ME overall for units with normed time span for OE: common bases from 25 to 400 clauses

The most substantial growth in SVO is evident in the two northernmost zones, with the rate decreasing in the other, more southern dialect sectors. The highest increase fluctuates, similar to the other output configurations, between 70 and 85 per cent. Furthermore, the layout generated for the higher common bases creates a familiar boost in the growth pattern in the Kentish sector. As before, this is a result of the loss of representativeness in the sample on the Old English side. The only noticeable difference between the current and the previous configurations is the sharper increase in the preference for SVO in what used to be an Anglian zone. On the whole, however, the pattern of the emergence of SVO remains comparable regardless of the modification made to the sampling units.

Finally, an attempt was made to examine the distribution of SVO when the material did not include texts which were translations from foreign originals (Fig. 6.1.1e). As a result, a substantial cut was made not only with respect to the number of texts available in sets but also to their size. As regards the number of texts, the exclusion of translations eliminated the Kentish unit. The material from the ME part of that unit contains texts that were all translations. When it comes to the text size, the OE material upon extraction of translated specimens was limited to manuscripts of a rather moderate size, hence the normalising to the lower common bases of 25 and 50 clauses.

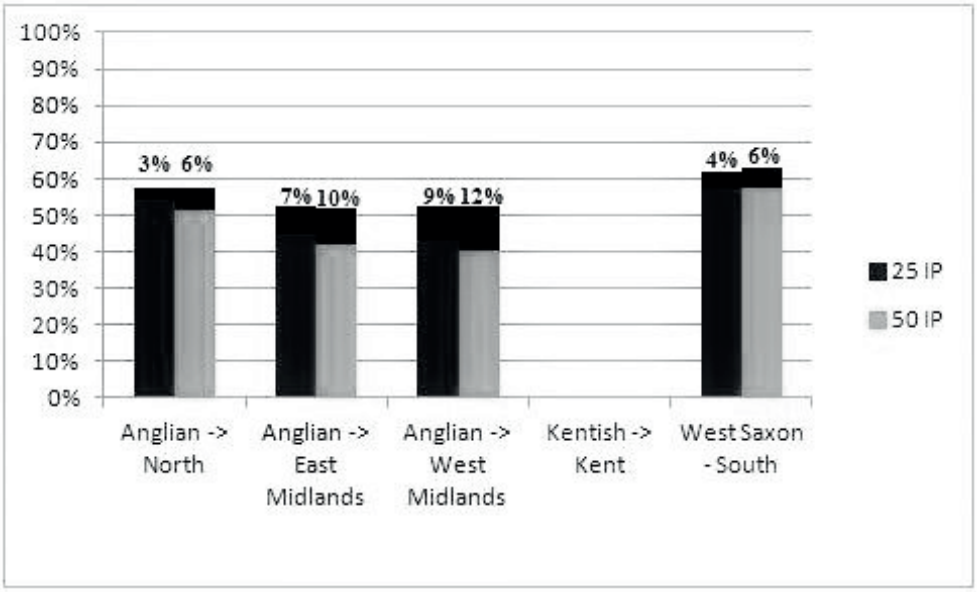

Figure 6.1.1e: Growth of SVO from OE to ME overall for units without translation: common bases 25 and 50 clauses

As indicated by the output columns in figure 6.1.1e, the content of sampling units with the least interference from translations clearly changes the distribution pattern of SVO. We can also notice that the differences between dialects have levelled out. However, with such dramatic alterations to the content of the OE database, one cannot provide a satisfactory explanation for the extent of the impact of translations diachronically. The distinct, levelled pattern may thus be nothing more than a result of the radical reduction of data.

Summing up the results of the comparison of the outputs for the OE and ME datasets, a few important observations have been made. Firstly, moving from the earliest days up to the end of the ME period, the rate of SVO increases, as could be expected. Secondly, each of the clause levels, matrix and subordinate, exhibits its own growth rate of the feature, which also confirms the commonly observed discrepancy between the two environments on that account (section 4.3.2; also Davis 2006: 52 and 103-4). The matrix clauses appear to have a more uniform distribution of SVO than the subordinate clauses. It is most likely a result of the fact that the former developed an SVO grammar, whether with or without V2, already during the OE period. By contrast, subordinate clauses did not develop a preference for SVO until the later stages, when other, not only language-internal factors, conditioned the growing prevalence of the feature. Thirdly, just as the clause types have their own rates of distribution, so do the sampling units (i.e. dialect sectors). Considering the fact that the distribution of the SVO sequence, as defined in the present study, is much more uniform within ME (cf. the following sections), the irregular growth pattern OE-to-ME has, in fact, confirmed how much the OE dialects differ from each other syntactically (cf. Corrie 2006: 86). Fourthly, the normalising of frequencies to varying common bases reveals shifts in the growth of preference for SVO. The rift usually occurs at the point of normalising

frequencies to 200 clauses or higher. Fifthly and lastly, while modifying the units to achieve a more homogeneous allocation of texts with respect to MS dating does not change the pattern of distribution of SVO, the exclusion of translated material alters the layout considerably. Still, it is difficult to assess whether the change was introduced genuinely by eliminating the interference from foreign originals or whether it was a shift induced by a drastic cut made on the content, especially that of the OE data.

6.1.2 The new syntactic layout in Middle English: dialectal variation in PPCME2

The distribution of SVO during the Middle English period, indeed, indicates a turning point, as far as syntactic conditions are concerned. The growth in the preference of SVO is substantial within all the sampling units, as indicated by the results presented in the previous section. At the same time, the level of variation seems to decrease, indicating that the feature in question was more established in Middle English than a few centuries earlier, as will be seen from the results below. SVO with a single finite verb constituted the majority of retrieved sequences. A few examples are presented below:

(10) Our Lord herd þe desire of þe pouer in gost; [ID CMEARLPS,11.403]
(11) Þis hope conforted me in my meknes, [ID CMEARLPS,148.6536]
(12) (…) þat he turned watyr ynto wyne, [ID CMMIRK,52.1469].

When compared with the output generated for the Old English database, the differences in average frequency between particular dialect sets are less prominent. Nonetheless, it is clear from the results table (see 6.1.2a) that the feature is present more in the northernmost dialects and in the south than in southeastern Kent or the West Midlands. The differences in frequency between the dialects grow with the increase of the common base. However, the contest for the highest average frequency values between the North and the South remains the same even when frequencies are normed to the common base of four hundred clauses. Furthermore, the average values seem to differ more between sets for the output that includes empty categories in the SVO order sequencing. On the whole, relying solely on the average frequencies does not offer enough valid input apart from the overall recorded increase in the preference of the feature for the entire period, a fact which conforms to the findings on the distribution of SVO generated so far.

Table 6.1.2a: Average frequency for SVO per dialect set overall for units including and excluding empty categories: common bases from 25 to 400clauses

SEQUENCING	INCLUDING EMPTY CATEGORIES			EXCLUDING EMPTY CATEGORIES		
F/COMMON BASE	25 IP	100 IP	400 IP	25 IP	100 IP	400 IP
N	7.18	28.72	116.40	5.56	22.22	89.75
EM	6.93	27.72	110.87	5.24	20.95	83.79
WM	6.29	25.16	100.66	4.61	18.43	73.73
K	5.95	23.81	95.23	4.62	18.48	73.91
S	7.48	29.94	119.74	5.60	22.39	89.56

$p < 0.0001$ for both comparisons within sequences including and excluding empty categories
$p < 0.05$ for comparisons between sequences including and excluding empty categories

More definitive conclusions can be drawn by observing the level of dispersion, as indicated by the values of the coefficient of variation in table 5.1.2b. Of the five distinct dialects, two show considerable regularity in the distribution of the SVO order for the sequencing including empty categories – the North and the South. More interestingly, however, only one repeats the same pattern for the SVO string which excludes the empty categories, viz. the northern set. This is a clear indication that the feature in question was well grounded there, bearing in mind that the frequency for this sector is one of the highest in the entire database.

Table 6.1.2b: Coefficient of variation per dialect set overall for units including and excluding traces and other empty categories: common bases from 25 to 400clauses

CV (%) / COMMON BASE	25 – 400 IP EXCL. EMPTY CAT.	25 - 400 IP INCL. EMPTY CAT.
N	6	6
EM	35	29
WM	34	31
K	15	16
S	8	20

The significance of this finding lies not only in the fact that the northern set corresponds to the area of external influence (Scandinavian and earlier Celtic) but also that it is the very same dialect that has been attested to first get rid of inflectional endings. As will be shown in due course, the pattern of high average frequency and substantial regularity (low CV) for the North is not a coincidence. The distinctly high preference for SVO will also be observed for particular clausal environments (especially at subordinate clause level), recorded both in the North and the East Midland sectors, which are the most Scandinavian influenced dialects. A noticeable increase in the regularity will be distinguished for configurations which exclude translations.

6.1.2.1 SVO at clause levels: matrix and subordinate

The results generated on the entire ME database, though enlightening with respect to assessing the growth in the preference for SVO during the period and pointing to interesting dialectal variation in that respect, do not provide the complete picture. The study of word order change in early English needs to be directed specifically at particular clause levels. Differentiation into matrix clause and subordinate clause environments in OE meant different word order preferences as well as different underlying rules which promoted these preferences, as presented in sections 4.3.2 and 5.2.2.

Matrix clauses readily exhibited the SVO order in the OE period, though other combinations were possible, largely to the presence of inflections (e.g. Irvine 2006: 47). Yet, a prominent constraint on the ordering flexibility was determined by the placement of the finite verb, which very often (though not always) was found in the second position in the clause. That is where OE was at its closest to other Germanic languages, which observed (and still do) the V2 rule within the matrix clause practically without exception (e.g. Fischer *et al.* 2000: 82). That said, matrix clauses could also exhibit V3 order, with pronominal subjects, as remarked in section 4.3.1.1.

By contrast, the word order of subordinate clauses in OE was predominantly verb-final (e.g. Fischer *et al.* 2000: 105 and 182, Swan 1994: 234-5). In addition, the frequent use of the subjunctive strengthened the syntactic marking typical of these types of clauses (e.g. Fischer *et al.* 2000: 89). Apart from different rules regulating the placement of finite verbs, subordinate clauses did not exhibit subject-verb inversion so freely as matrix clauses did (Haeberli 2001: 215). By the time the Middle English period ended, word order ceased to play a crucial role in differentiating clause types. However, the clausal asymmetry was still visible especially during early ME, even though subordinate clauses no longer showed verb-finality so overtly as during the OE times (e.g. Haeberli 2001: 225).

The results generated from the PPCME2 material have been arranged by dialect and the element sequencing (Table 6.1.2.1a and b), i.e. the SVO order which included or excluded empty categories (ECs). Normally, the sequences which exclude the empty elements have been essentially considered in the present study. The composite nature of the EC label does not allow any definitive conclusions regarding their individual input in the sequences, as mentioned in section 5.2.2. However, the inclusion of the contrasting sequencing option has been added here for comparison. The distribution of SVO which includes ECs allows, among others, assessing whether a particular layout of SVO, especially with respect to the distinct clause levels and dialect sectors is random. In the same vein, the method of multi-common-base comparisons has been observed in the current research to address the issue of different text lengths in the sectors.

As table 6.1.2.1a shows, the average normed frequencies observed for particular dialects at the matrix clause level oscillate within a similar range, with the smallest difference for a given common base amounting to only 0.05 and the biggest one reaching 16 instances. Interestingly enough, chi-square tests indicate that the differences between particular dialects are statistically significant in each comparison round. Still, the highest recorded values for a given common base do not form a uniform pattern. More often than not, it is the case of the North vs. the South with respect to the most overt preference for SVO. The lack of a clear pattern when recording the highest normed frequency is not limited to one particular sequencing option but occurs each time a common base is changed, whether the SVO string includes or excludes ECs. Of the two configurations, the latter shows more potential for uniformity. The discrepancies between the sequencing options suggest that the rules governing the use of empty categories might constitute yet another characteristic in determining the frequency of occurrence of the investigated feature.

Table 6.1.2.1a: Average frequency for SVO per dialect set at matrix clause level for units including and excluding empty categories: common bases from 25 to 200 clauses

SEQUENCING	INCLUDING EMPTY CATEGORIES			EXCLUDING EMPTY CATEGORIES		
F/COMMON BASE	25 IP	100 IP	200 IP	25 IP	100 IP	200 IP
N	8.56	32.99	60.05	7.98	31.28	54.87
EM	7.78	31.14	62.28	6.68	26.71	53.42
WM	8.10	32.40	64.80	6.73	26.90	53.81
K	7.42	29.66	57.10	6.94	27.76	54.62
S	9.18	26.74	73.54	7.55	30.19	57.19

$p < 0.0001$ for all comparison options

In addition, the highest normed frequency for SVO recorded in the North and the South could likely be the effect of the uneven distribution of the earlier texts within the dialect sets. For both the North and the South, the first part of the Middle English period is virtually not represented and the high frequency recorded in these sets within the matrix clause environment might be the result of calculations made on the material coming from the second part of the linguistic epoch. Still, the data in the northern set seems to be more auspicious on that account. The dates of composition of some of the texts from this dialect are different from the dates of the manuscripts. The former point to a much earlier period, with the latter designating dates of copies, thus suggesting the continued use and existence of earlier linguistic habits in later reproductions[127] (e.g. Corrie 2006: 102). As regards misrepresentation in terms of subperiods in PPCME2, the Kentish dialect, too, is problematic. In this case, the shortage of data concerns the later period of ME.

By contrast, the situation within the subordinate clause environment (Table 6.1.2.1b) is quite transparent. Although the normed frequencies again vary within a similar range, a consistent pattern can be distinguished in the highest average normed frequencies recorded in the sectors. The northern dialect is where the most overt preference for SVO can be observed throughout the configurations, viz. in all the common base rounds and for both the sequencing options. The second highest frequency is mostly found in the East Midlands. This regional preference for SVO at the subordinate clause level is indeed striking. This deserves some attention not only because of the contrasting irregularities in the distribution found at the matrix clause level but also because the differences in the average frequencies between the sectors were not very definite for the majority of configurations generated in the present study.

[127] I.e. copies preserving the linguistic features of its exemplars.

Table 6.1.2.1b: Average frequency for SVO per dialect set at subordinate clause level for units including and excluding empty categories: common bases from 25 to 200 clauses

SEQUENCING	INCLUDING EMPTY CATEGORIES			EXCLUDING EMPTY CATEGORIES		
F/COMMON BASE	25 IP	100 IP	200 IP	25 IP	100 IP	200 IP
N	6.46	25.86	54.72	4.27	17.08	35.99
EM	6.16	24.65	49.29	3.72	14.86	29.73
WM	5.20	20.82	41.63	3.25	13.00	26.00
K	4.93	19.72	39.44	3.01	12.03	24.05
S	5.92	23.70	47.40	3.72	14.88	29.77

$p < 0.0001$ for all comparison options

A comparison between the results in tables 6.1.2.1a and 6.1.2.1b reveals a noticeable difference between the two clause levels in the average normed frequencies of SVO. The two clausal environments are indeed distinct in their distribution of the SVO tokens, as indicated by the p-values calculated from the chi-squares[128], which tend to stay below 0.0001, pointing to differences between the two as being (extremely) statistically significant. Expectedly, the matrix clause environment clearly manifests SVO grammar, at least partially due to the lingering on of old syntactic patterning. The subordinate clauses, on the other hand, were only beginning to show the effects of the structural makeover. The average normed frequencies recorded in particular dialect sets are thus not as high. Still, considering the fact that the ordering of clausal components within the subordinate clause environment had been so different in OE and regulated by completely distinct mechanisms, the presence of SVO at this clause level during ME and its proportion to the rate of the feature recorded within the matrix clause environment appear to be rather congruent. This raises the question whether some of the processes within one level could affect the other. On one hand, there is a proposal by Stockwell and Minkova (1991), who claim that this may have been the case. According to them, some of the subordinate clause structural effects[129], notably the subject as a topic creating SV, could have been mapped from the matrix clause environment on to the subordinate one (1991: 394). On the other hand, there is Haeberli (2001: 227) who claims that the processes observed in main clauses could not occur in subordinate clauses. He follows Bobaljik and Thrainsson's (1998) proposal which stresses the main/subordinate asymmetries[130].

Another question to ask when observing the growing prevalence of SVO at IP-SUB is why the distribution of the feature is so dialectally conditioned for this clause level. Not only does it align with the areas considered to be the bases of the onset of inflectional erosion, but it also corresponds to the regions affected by the early language contacts. Two contact scenarios need to be considered at this point: the impact from the Celtic and the Scandinavian languages. Unlike the influences from Latin and French, these were geographically specific phenomena (e.g. Corrie 2006: 97). While

[128] Each for a given sequencing option, since both return a different number of tokens.
[129] Stockwell and Minkova refer to them as "the subset of main-clause verb-second exemplars" (1991: 394).
[130] The assumption here is that finite verbs do not move to the same position in the two clause types (Haeberli 2001: 227).

the northern dialect corresponds to the areas affected by the contacts with both the Celtic and the Scandinavian populations, the East Midlands dialect belongs exclusively to the Norse zone (e.g. White 2002: 154). Naturally, one cannot completely rule out the existence of pressure coming from the Celtic source, though its input might have already been diluted with time. The repeated, more prominent preference for SVO particularly in the North, recorded for the total of IP-MAT and IP-SUB instances, can be indicative of the working of both influences. Yet, when additionally analysing the regularity of the distribution of SVO at the subordinate level, one is bound to notice that the two dialect sectors coalesce not only with respect to the average normed frequencies but also in their values of the coefficient of variation (CV) shown in Table 6.1.2.1c below. This tips the scales, again, more in favour of the Scandinavian influence, where the two dialect sectors are perceived as a single, contact-induced area.

Table 6.1.2.1c: Level of variation for SVO per dialect set at subordinate clause level for units excluding and including empty categories: common base from 25 to 200 clauses

SEQUENCING	**INCLUDING EMPTY CATEGORIES**			**EXCLUDING EMPTY CATEGORIES**		
CV/COMMON BASE	25 IP	100 IP	200 IP	25 IP	100 IP	200 IP
N	16	16	8	26	26	24
EM	33	33	33	25	25	25
WM	43	43	43	53	53	53
K	21	21	21	31	31	31
S	24	24	24	10	10	10

As shown by the tables, the similarity of the CV values can also be found in the southern dialect pair (i.e. Kent and the South), occurring at the matrix clause level (Table 6.1.2.1d), too, and with the sequences including the empty categories. It is possible, therefore, that these parallel values are simply the result of a particular sequencing option (i.e. coincidental) rather than evidence of the effect of substantial interference from other languages. Other factors might be at play as well, especially if the generally acknowledged contrast between the more innovating northern and East Midland and the more conservative[131] Kentish and southern dialects is taken into account (e.g. Millward 1996: 142, Corrie 2006: 95-8). Nevertheless, it is intriguing to find this coalescence only for the northern and/or the southern pair, with the output of the West Midland dialect always cutting through the two and practically never[132] lining up with either.

[131] The relevance of the contrast applies strictly to morphosyntax. It is common knowledge that the mechanisms regulating some of the phonological tendencies of the Middle English period generated the difference between the sectors in the opposite direction (e.g. Millward 1996: 142).

[132] The CV values for West Midlands approach those displayed for the South at the matrix level for the sequences excluding the empty elements.

Table 6.1.2.1d: Level of variation for SVO per dialect set at matrix clause level for units excluding and including empty categories: common bases from 25 to 200 clauses

SEQUENCING	INCLUDING EMPTY CATEGORIES			EXCLUDING EMPTY CATEGORIES		
CV/COMMON BASE	25 IP	100 IP	200 IP	25 IP	100 IP	200 IP
N	21	23	13	26	29	10
EM	27	27	27	38	38	38
WM	9	9	9	14	14	14
K	10	10	12	9	9	12
S	14	14	17	16	16	14

The incongruity of the West Midland sector with respect to the level of variation at both clause levels for both sequencing options is, indeed, peculiar. The lack of common ground with the northern pair or even exclusively with the East Midland set could be explained, at least, by the effects of the contact with Norse (or lack of it) (Corrie 2006: 91). In addition, the contrast might also be indicative of the presence of 'old' versus 'new' with respect to the maintenance of the early OE literary tradition that the (south) West Midland school was famous for (Corrie 2006: 109). However, it is hard to explain why the West Midlands, for the most part, does not align with the southern pair as well. The more conservative manner of expressing morphosyntactic relations thrived there; yet no overt (i.e. repeated) similarity in the values of the CV was recorded. It is conceivable that the effects of socio-linguistic changes from 1350 onwards did not percolate through the sectors in the same manner. It is also possible that the typical intermediary character of the West Midland dialect manifested itself most clearly at that point (e.g. Brook 1972: 68). Since both the East Midland and southern features were present in this dialect, the distinct combination of these various tendencies would make the WM set stand out on its own.

So far, we have climbed merely one step higher on the scale of structural specificity – going from the all-encompassing sentence level to particular clausal environments. It has been shown that some of the tendencies recorded in the former are also present in the latter, for example, in the distinct preference for SVO in the North. It is important to examine whether the patterns recorded up to now occur with other configurations, when other factors known to influence the distribution of the feature are explored. The following sections deal with the impact of these factors, notably the interference from foreign translations, particular genres as well as some aspects of MS dating.

6.1.2.2 Early and later material: subperiod comparisons

Middle English has often been described as a period of great changes caught in the whirl of more or less local linguistic tendencies (e.g. Horobin and Smith 2002: 52). Many of these changes and tendencies were in direct relation to the socio-political situation of the post-Conquest England. As the present study has aimed at assessing the role of inflectional erosion (notably the loss of cases) in the subsequent emergence of the rigid SVO word order, the timeline concerning the establishment of the syntactic components, pertaining to that particular configuration, is of essence. Furthermore, since the Middle Ages in England appear to be divided, both with respect to social

politics and language politics, into an earlier and later phase, it is vital to determine if and how much the great changes in ME currently discussed correlate with the events and practices of these two subperiods. The dividing line between these phases can confidently be set in the 1350s. With respect to the chronology of the morphological simplification and emergence of SVO, it appears that there could be a nexus between the two. When one process reaches the end of the line, the other begins to emerge. Clear signs of disintegrating inflectional morphology could already be detected in the texts at the end of the OE era. The culmination of this change, nevertheless, took place in the first half of the ME period, with the northern dialects adopting the change earlier than other linguistic regions. The distinct case endings on nouns, responsible e.g. for marking of the subject and object in a clause, petered out completely by the 13th century (e.g. Lightfoot 1991: 123). With respect to the growing prevalence of SVO, it has often been indicated that this syntactic configuration was dominant from the beginning of ME (e.g. Fischer *et al.* 2000: 206). Some scholars, notably Swan (1994), point to the word order being "fairly well established" already in the later OE (Swan 1994: 235). As will be shown in the results tables below, this early attested volume of popularity of SVO was a rather mild sign of the things to come, particularly with respect to the level of entrenchment of the feature.

The results presented below have been generated on the same samples as the output from the previous sections. This time, the corpus material was divided into two sets, representing early ME (1150-1350) and late ME (1350-1500). Raw frequency counts were normed to three distinct common bases. The highest common base was set at 300 clauses (IP) in order to assess whether the contrast between the earlier and later distributions deepens or weakens in particular dialect sets. In addition, as mentioned previously, certain subperiods are not covered text-wise in three out of five dialects of the epoch. The North and the South lack texts from the first half of the period while Kent is excluded from comparison for the second part of ME. That being said, calculations were made for the northern set during the period of 1150-1350. The material from this section included texts with an earlier date of composition and a later date of manuscript. If one accepts the idea of earlier linguistic features lingering on in later reproductions (e.g. Corrie 2006: 102), the figures in grey/faded font may be considered. Finally, much as in the case of IP-MAT vs. IP-SUB distributions, the sequencing option which incorporates empty categories has been made available to ensure that the retrieved patterns are not random.

Even a cursory glance at the normed frequencies for both sequencing options (tables 6.1.2.2a and b) reveals the expected difference between the sets representing the earlier and later subperiods of ME. At almost all times, there is an increase in the preference for SVO during the second half of the epoch. The increase is more prominent for the SVO string without the empty elements.

Table 6.1.2.2a: Average frequency per dialect set overall for units including empty categories. Comparison between first and second half of ME: common bases from 25 to 300 clauses

F	SVO during 1150-1350 25 IP	SVO during 1350-1500 25 IP	SVO during 1150-1350 100 IP	SVO during 1350-1500 100 IP	SVO during 1150-1350 300 IP	SVO during 1350-1500 300 IP
N	7.39	7.07	29.56	28.30	88.67	86.38
EM	6.87	6.99	27.49	27.95	82.47	83.84
WM	5.08	7.90	20.33	31.61	60.99	94.83
K	5.95	n/a	23.81	n/a	71.43	n/a
S	n/a	7.48	n/a	29.94	n/a	89.81

$p < 0.0001$ for all comparison options

Table 6.1.2.2b: Average frequency per dialect set overall for units excluding empty categories. Comparison between first and second half of ME: common bases from 25 to 300 clauses

F	SVO during 1150-1350 25 IP	SVO during 1350-1500 25 IP	SVO during 1150-1350 100 IP	SVO during 1350-1500 100 IP	SVO during 1150-1350 300 IP	SVO during 1350-1500 300 IP
N	5.44	5.61	21.77	22.45	65.31	68.65
EM	4.86	5.61	19.44	22.45	58.32	67.36
WM	3.69	5.84	14.74	23.36	44.22	70.07
K	4.62	n/a	18.48	n/a	55.43	n/a
S	n/a	5.60	n/a	22.39	n/a	67.17

$p < 0.0001$ for all comparison options

Looking at the earlier part of ME alone reveals yet another intriguing consistency. Of the three dialects available for comparison, viz. the East and West Midlands along with Kent, the most SVO friendly set is without exception the East Midland one. There could be a couple of explanations for this state of affairs. First, the clear preference for SVO within the East Midland sector at such an early stage points to the high level of distribution of the feature already during the onset of the process of the spread of innovations. This process entailed the emergence of new syntactic preferences, as concomitant with the inflectional erosion which was largely completed during the first half of ME. If the scenario of the progression of change at issue north-southwards (e.g. Tristram 2002: 125) is accepted, the average normed frequency values recorded for the East Midland set would make the innovative North - the starting point of progression - the location where the prevalence of SVO was already well established during these early times. Naturally, one might wonder why the distribution of SVO was not equally popular in other areas of the Midlands (West). To account for this, one should bear in mind that the East Midland dialect belonged to the areas of early Scandinavian influence. The link between the North and the East Midlands has already been asserted in the previous section. It is only logical to see the East Midland set as the area where

the loss of (case) inflections accompanied by the growing prevalence for SVO was accepted so early during the Middle English period. Further, the second highest average normed frequency for the data from the first half of ME should be found in the West Midland set. The fact that the Kentish dialect takes over with the preference for SVO for the sequences including the empty categories could be ascribed to the presence of the categories themselves.

As to the second part of the Middle English period, a more pronounced preference for SVO, indeed, coincides with the dynamic socio-political situation in England and the changing attitudes towards the English language at that time. The date of 1350 marks the new beginning with respect to the restoration of both the English identity and language. With the official adoption of English as the language of the State and of instruction more and more written material was available to the public (section 2.4). At the same time, writers and copyists were encouraged to use and celebrate their local vernaculars (e.g. Corrie 2006: 102). Analysing the distribution of the feature on the data originating during these two (and a half) centuries constitutes the most valid picture of linguistic preferences as plotted on what seems to be the most accurate dialectal map of medieval English.

A complementary piece of evidence with respect to estimating the volume of contrast in the distribution of SVO before and after the 1350s is provided by observing the level of variation shown in table 6.1.2.2c. Admittedly, there are only two dialect sectors to supply the information on that account – the East and West Midlands. Only these two sets contain ample material from the entire Middle English period. Nevertheless, the values of the coefficient of variation (CV) recorded in the two sectors are distinct enough for comparison. In addition, it has been shown so far that although geographically adjacent, the two sets stand apart in their preference for this feature whether at the sentence level or within particular clause types.

Table 6.1.2.2c: Coefficient of variation per dialect set overall for units excluding and including empty categories. Comparison between first and second half of ME: common bases from 25 to 300 clauses

	EXCLUDING EMPTY CATEGORIES		**INCLUDING EMPTY CATEGORIES**	
CV	SVO during 1150-1350 25,100, 300 IP	SVO during 1350-1500 25,100, (300) IP	SVO during 1150-1350 25,100, 300 IP	SVO during 1350-1500 25,100, 300 IP
N	6	7(6)	5	7
EM	52	23	40	24
WM	35	15	32	10
K	15	n/a	16	n/a
S	n/a	8	n/a	20

The degree of prevalence of SVO suggests that this pattern was well established until the latter part of the era. The CV values decrease by a half in the EM and WM dialect sets during the period of 1350 to 1500 for all common bases adopted and for both

sequencing options. Furthermore, the low CV obtained for the North and South during later ME also indicates that SVO was becoming more prevalent. Of the two sets, the North is the most consistent in showing the regularity of distribution of SVO, as the values of the coefficient remain the lowest even for the results from the searches on sequences which include empty categories (columns on the right). It needs to be emphasised that CV values below 10 per cent signal a particularly high degree of uniformity among the texts of a given dialect with respect to the preference of the feature.

The conventional view holds that the structural changes between the end of the OE period until around the 14th century of ME would have occurred one way or the other, and the impact of either ON[133] or, later on, of the Norman language would have been of no consequence. Still, it is at least striking that the material from the areas affected by contacts with the Scandinavian population is more advanced with respect to the changes which soon were to prevail than the text specimens belonging to the dialects which were not exposed to these external pressures. As for the second half of ME, a clear rise in the preference for SVO was recorded, reaching all the regions on the dialectal map of the period.

6.1.2.3 Interference from translations

For efforts to track the pattern of the emergence of the SVO order in Middle English, the presence of translations in PPCME2 inevitably introduces a distorting factor. The distinct syntactic input offered by foreign originals at that time, whether Latin, (Old) French or even Dutch, needs to be taken into account (e.g. Millward 1996: 187). Not only did the reliance on translations mean the possibility of bringing in foreign elements (usually vocabulary) but also, quite often, it introduced straight-forward calquing of entire structures, which were not exactly on a par with how the target language would normally behave (e.g. Davis 2006: 127). In order to estimate the level of interference from translations, the results of the distribution of SVO generated on the largest, unrestricted samples had to be compared with the output which did not contain translations. The samples were, thus, arranged to include the least amount of text influenced by foreign originals.

Subtraction of the translated material, quite predictably, does not create major fluctuations in the average normed frequencies between the sampling units (Table 6.1.2.3a). On the whole, there is a slight decrease recorded in all sectors for both sequencing options, once the potential foreign impact is taken out of the equation, with the exception of the East Midland dialect at the subordinate level, for sequences excluding empty categories. In this particular case, the sector displays an increase in the preference for SVO, making the Scandinavian zone (the North and the East Midlands) return the highest values for the feature in question. The tendency remains unchanged in the output calculated for the higher common bases, although the differences between the dialects get more pronounced[134]. Further, the shift in normed frequencies in IP-MAT between the two sequencing options seems to be created precisely by the inclusion/exclusion of empty categories.

---

[133] Also Celtic and Latin for that matter.
[134] Tables for higher common bases are presented in Appendix 6.

Table 6.1.2.3a: Average frequency per dialect set at all clause levels, both sequencing options. Comparison between the content of the entire ME period and the material excluding translations: common base of 25 clauses

F/25 IP	SVO IP-	SVO IP- no transl.	SVO IP-MAT	SVO IP-MAT no transl.	SVO IP-SUB	SVO IP-SUB no transl.
Sequencing including empty categories						
N	7.18	7.01	8.56	8.23	6.46	6.27
EM	6.93	6.28	7.78	7.01	6.16	5.42
WM	6.29	6.30	8.10	8.22	5.20	5.29
K	5.96	n/a	7.42	n/a	4.93	n/a
S	7.48	7.16	9.18	9.10	5.92	5.62
Sequencing excluding empty categories						
N	5.56	5.53	7.98	7.44	4.27	4.37
EM	5.24	4.62	6.68	5.55	3.72	3.89
WM	4.61	4.47	6.73	6.28	3.25	3.44
K	4.62	n/a	6.94	n/a	3.01	n/a
S	5.60	5.58	7.55	7.81	3.72	3.63

IP-MAT comparison between sequences with and without the empty categories $p < 0.05$
Other comparison options $p < 0.0001$

The distinctive character of the Scandinavian influenced regions becomes more prominent when the level of variation is added to the picture (Tables 6.1.2.3b and c). Upon exclusions of translations for the sequencing including empty elements, the majority of sampling units show an increase in variation in the distribution of SVO. The East Midland set, however, reveals exactly the opposite tendency regardless of the clause level. The biggest leap towards the more regular distribution is revealed, again, at the subordinate clause level.

Table 6.1.2.3b: Coefficient of variation per dialect set at all clause levels for units including empty categories. Comparison between the content of the entire ME period and the material excluding translations: common base of 25 clauses

CV/25 IP	SVO IP-	SVO IP- no transl.	SVO IP-MAT	SVO IP-MAT no transl.	SVO IP-SUB	SVO IP-SUB no transl.
N	6	8	21	24	16	23
EM	29	21	27	26	33	14
WM	31	41	9	11	43	52
K	16	n/a	10	n/a	21	n/a
S	20	26	14	20	24	33

For sequences excluding empty elements (Table 6.1.2.3c), the growth in regularity is additionally extended to the northern set and, in one case (matrix clauses), for the West Midlands, when leaving out translations. Nevertheless, there is only a slight drop in variation in the latter, whereas both the North and the East Midlands show quite a substantial increase in the uniform distribution (i.e. low CV values) of SVO.

Table 6.1.2.3c: Coefficient of variation per dialect set – all clause levels for units excluding empty categories. Comparison between the content of the entire ME period and the material excluding translations: common base of 25 clauses

CV/25 IP	SVO IP-	SVO IP- no transl.	SVO IP-MAT	SVO IP-MAT no transl.	SVO IP-SUB	SVO IP-SUB no transl.
N	6	4	36	19	26	27
EM	35	31	38	47	25	10
WM	34	46	14	13	53	63
K	15	n/a	9	n/a	31	n/a
S	8	11	16	20	10	13

In order to ensure that the drop in variation within the northernmost sectors is not a result of a random cut based on the larger unrestricted samples, a repeated sampling procedure was conducted on the large amount of material available in the East Midland dialect. As many as twenty-one (A-U) distinct samples were created with the level of variation compared before and after the subtraction of translations (Table 6.1.2.3d below).

Table 6.1.2.3d: Coefficient of variation - Repeated sampling on the EM material with units including and excluding translations: Sequencing without empty categories – overall (total of IP-MAT and IP-SUB)

	STAND.	A	B	C	D	E	F	G	H	I	J
Incl. transl	35	35	32	35	31	31	35	33	32	31	35
Excl. transl	31	38	27	38	28	28	31	28	26	28	21
	K	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**
Incl. transl	34	34	34	39	34	35	25	34	35	34	37
Excl. transl	28	28	24	47	30	11	28	27	12	29	10

$p < 0.0001$ for all comparison options

For the option excluding empty elements from the SVO sequences, the decrease in variation (i.e. the increase in regularity) was recorded in 17 samples, that is, in over 80 per cent of cases. At this point, it is clear that the interference resulting from the presence of translated material within sampling units has to be accounted for. What is more, the exclusion of translations revealed the tendency which would prevail predominantly in the sectors corresponding to the regions of Scandinavian influence. The contrast between what would be recognised as syntactically consistent English and the content created in the structural fashion of foreign originals is most distinguishable in these two dialects. Further, the impact of translations seems to conceal a recurring tendency of high preference for SVO at the subordinate level – a development that is also characteristic of the dialect sectors where Scandinavian presence was once felt most.

6.1.2.4 Impact of genres and other factors influencing distribution of SVO

Apart from the information on the date of composition/manuscript and on whether a given specimen is a translation from a foreign original, the PPCME2 material has been classified according to genres. Not only has it been a useful point of reference considering the data available but also a necessary aspect to take into account when analysing the distribution of word order, whether synchronically or diachronically. Just as the interference from foreign translations induces changes in the distribution of the feature so does the impact of text type, as emphasised in section 5.2.2.

In order to eliminate as much as possible the lack of symmetry between various genres, samples were also arranged to represent material from (larger) prototypical text categories (PTC) (section 5.3.4.3). Table 6.1.2.4a below presents the results of comparisons between the distribution of SVO within the dialect of the East Midlands. The substantial number of texts available from this dialect allowed complex comparison between the standard samples[135], the entire 'population'[136], that of particular PTC, and finally, a single genre within one of the available PTCs. Three distinct PTC samples were created: secular instruction (IS), religious instruction (IR), and non-imaginative narration (NN). The last sample contains solely examples of religious treatise, which belongs to IR.

Table 6.1.2.4a: Average frequency and coefficient values for the East Midland dialect, both sequencing options. 'Population', sample and genre group (PTC) comparisons: common base of 25 clauses

Sequencing option	**Frequency** / **Coefficient of variation**	**POPULATION 1150-1500**	**STANDARD SAMPLE 1150-1500**	**PTC: religious instruction (IR)**	**PTC: non-imaginative narration (NN)**	**PTC: secular instruction (IS)**	**SINGLE GENRE WITHIN IR: religious treatise**
SVO without the empty categories	EM F/25	5.38	5.24	5.11	5.75	5.35	5.10
	EM CV	25%	35%	27 %	30%	15%	22%
SVO including the empty categories	EM F/25	6.63	6.93	6.22	7.47	6.40	6.35
	EM CV	21%	29%	22%	17%	15%	22%
Number of texts		25	6	13	4	5	8

all comparison options $p < 0.0001$

Interference from genres seems detectable though minor. The differences are found more in the coefficient values than in the average (normed) frequencies. The value of the CV ebbs slightly from the prevalent pattern (lower values) in three out of six sequencing possibilities for genre group sets – NN for SVO including empty categories and IS for SVO in both sequencing options. The former (NN) also reveals a slight

[135] Samples used in the majority of comparisons in the present study.
[136] All the texts available in the EM dialect. Both standard and population samples contain texts which belong to different genres.

departure from the norm in the average frequency of the configuration including empty categories. Taking into account the fact that the results calculated for NN where sequencing includes the empty elements differ from the prevailing pattern both in frequency and CV values, it is possible that the difference is due to the inclusion of the empty categories alone. The rest of the output conforms both to the values calculated for the entire population as well as the standard sample used in the current study. Moreover, the division into larger prototypical categories to control the genre variation between units is adequate but apparently causes no significant change in the distribution of SVO. We can notice, for instance, that the output generated for a single genre, religious treatise, conforms to that generated for the PTC category of religious instruction, which includes the genre in question. At the same time, the results from this single genre display no sharp deviations from the results calculated on the standard sample or on the entire 'population' of East Midlands.

Table 6.1.2.4a also reveals that the values of the standard sample quite accurately mirror the results recorded for the entire 'population', which does not seem coincidental. Once the dating factor is taken into account, with every subperiod represented, both the average frequencies and the values of the coefficient are alike. In order to show how much the represented subperiodisation matters, the charts below present the results of the repeated sampling procedure. Sixteen distinct samples were generated from all available texts in the East Midland dialect. These samples repeat the layout of the standard sample used in the current study, where all subperiods are present in the same manner (Tables 6.1.2.4b and c). The results for both sequencing options show that there is no serious departure from the pattern either in the average frequency values or in the level of variation. The East Midland set stays within the expected flux of 6.24 – 7.50 in frequency and 21-38 per cent in CV values.

Table 6.1.2.4b: Repeated sampling for the East Midlands: Uniform subperiodisation, sequences with empty categories: common base of 25 clauses

EM	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16
F/25	6.58	6.46	6.31	6.24	6.65	6.77	6.36	7.14	7.29	6.45	6.78	7.35	6.48	7.51	6.45	6.90
CV	32	32	34	31	31	29	33	23	21	30	26	22	29	23	27	26

A significant finding, however, as far as the stability of the units is concerned, is to be seen in the resampling results table generated on the SVO strings which exclude the empty categories (Table 6.1.2.4c below). While the values of the coefficient fluctuate mildly for the sequencing option incorporating empty elements, they are virtually identical when the empty categories are taken away from the samples.

Table 6.1.2.4c: Repeated sampling for the East Midlands: Uniform subperiodisation, sequences without empty categories: common base of 25 clauses

EM	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16
F/25	5.04	4.96	5.03	5.03	5.08	5.65	5.02	5.35	5.46	4.72	5.26	5.47	4.79	6.02	4.67	5.08
CV	35	34	36	32	35	36	34	32	31	36	31	31	35	35	34	34

$p < 0.0001$ for all comparisons

Not only does it show that the output in the units remains constant when subperiodisation is observed, but it also clearly illustrates that the division into the two sequencing options is meaningful. Moreover, the samples contain different files (i.e. different text types) in every sampling round, which seems to be indicative of the dating factor having a greater impact on the distribution of the SVO order than the influence coming from genres.

A rather insignificant impact of particular text types can also be distinguished when all the sampling units (dialect sectors) of the Middle English corpus are taken into consideration. The availability of sources representing a single PTC of religious instruction (IR) enabled creating an output chart including almost every dialect of PPCME2. The results generated for IR were set against units used as standard samples in the present research, which contained texts from several genres representing distinct PTCs.[137]

Table 6.1.2.4d: Comparison of standard samples with the sampling units containing texts from a single PTC, sequencing with and without empty categories: common base of 25 clauses

		Non-specified genre groups, SVO including empty categories	IR(religious instruction) group, SVO including empty categories	Non-specified genre groups, SVO excluding empty categories	IR(religious instruction) group, SVO excluding empty categories
North	F/25	7.18	7.43	5.56	5.60
	CV	6%	3%	6%	7%
East Midlands	F/25	6.93	5.96	5.24	4.63
	CV	29%	33%	35%	36%
West Midlands	F/25	6.29	5.81	4.61	4.15
	CV	31%	34%	34%	36%
Kent	F/25	5.95	n/a	4.62	n/a
	CV	16%	n/a	15%	n/a
South	F/25	7.48	8.29	5.60	5.94
	CV	20%	15%	8%	7%

all comparison options $p < 0.0001$

[137] Samples for the Kentish set originally contained texts which all belonged to the IR category. They were, thus, excluded from this comparison.

As can be seen in table 6.1.2.4d, the results calculated for the sampling units created on the material from a unified genre group (IR) are almost the same as the output drawn on the standard samples. This fact was already acknowledged in table 6.1.2.4a, where only data from the East Midland dialect were shown. It has to be noted that the 'IR round' (right column in each sequencing pair) was created by removing texts[138] from the standard round (left column) which did not belong to that prototypical text category. This resulted in some overlap in values, as could be expected. However, it looks as if virtually nothing had happened in the IR set after the extraction, even though some of the sets lost half of their original content. Not surprisingly, the normed frequencies vary somewhat but overall the proportions remained the same. The highest normed frequency is always found in the South. The North comes next, followed by the East and West Midlands, respectively. Most important of all, the CV values in both rounds vary within a similar range. This is a crucial finding. Typically, the coefficient of variation reacts noticeably to any changes in the content of the samples. The statistic could be expected to respond to a shortage of texts within sets if some other factor - in this case the dialectal type - did not strongly determine the distribution of the feature. Nonetheless, the table shows that CV remains relatively steady. It follows, therefore, that even if genres exert some influence on the distribution of SVO in particular dialects it seems not to reflect in any major way in the tables. Either the impact of genres does not override dialectal complexity or the material included in PPCME2 is not sufficient to comprehensively test the extent of this interference.

The fact that the power of dialect overrides the potential of text type, or any other influencing factor for that matter, can already be acknowledged when comparing the output from text doublets (i.e. dialect translations) (Diensberg 1996: 256) (Tables 6.1.2.4e and f).

Table 6.1.2.4e: Comparison of dialect translations, sequencing excluding empty elements: common base of 25 clauses

	IP-25 IP	IP-MAT 25 IP	IP-SUB 25 IP	IP-100 IP	IP-MAT 100 IP	IP-SUB 100 IP	IP-200 IP	IP-MAT 200 IP	IP-SUB 200 IP
cmedthor	6.12	6.84	5.57	24.48	27.35	22.28	48.96	54.70	44.57
cmedevern	4.78	8.07	4.03	22.96	32.28	16.11	45.93	64.57	32.21

$p < 0.003$

Table 6.1.2.4f: Comparison of dialect translations, sequencing including empty elements: common base of 25 clauses

	IP-25 IP	IP-MAT 25 IP	IP-SUB 25 IP	IP-100 IP	IP-MAT 100 IP	IP-SUB 100 IP	IP-200 IP	IP-MAT 200 IP	IP-SUB 200 IP
cmedthor	7.38	7.24	7.48	29.50	28.94	29.93	59.01	57.89	59.87
cmedevern	7.51	8.35	6.89	30.04	33.39	27.58	60.08	66.77	55.16

$p < 0.0001$

[138] The North lost 2 texts out of 6, the East Midland set, 3 out of 6, the West Midland sector, 2 out of 6 and the South, 3 out of 5. More details can be found in Appendix 8.

The files <cmedthor.m34> and <cmedevern.m3> contain the same text, *Mirror of St. Edmund*, representing the northern and the West Midlands dialects, respectively. The individual frequencies differ in the two versions. The dialectal disposition is visible within a majority of the sequencing options (e.g. MAT vs. SUB) and the common bases adopted. The northern version, predictably, exhibits a higher preference for SVO, with the subordinate environment particularly accepting the 'new' word order. The difference is smaller in the case of the SVO sequences which include empty categories. Still, even for this sequencing option the highest preference for the investigated feature does not diminish at the subordinate clause level. The preference for SVO at IP-SUB in the North was already recorded within standard samples (section 6.1.2.1).

The above observations on factors influencing the distribution of SVO in Middle English allow us to roughly rank them on a provisional scale of importance. Dialect background seems to override all other factors. With respect to frequency of occurrence, the analysis of dialect translations has shown that the differences between the two versions correspond to the distributional tendencies exhibited for standard samples and within several output configurations. The dominance of dialect can also be seen in the rather stable values of the coefficient of variation, which shows that the samples used in the current study are representative. To illustrate this, a closer look at the CV recorded for the East Midland dialect in the majority of the tables included in the current chapter reveals that the variation occurs mostly within the range between 20 to 38 per cent. A similar level of CV values is maintained whether a different subperiodisation in samples is used, whether translated texts are included in the sets and whether the data selection is limited to prototypical text categories (or even genres). At the same time, the range of variation found within the East Midland dialect needs to be contrasted with the very low CV values (3 up to 7 per cent) frequently found in the North and also in the South.

The impact of other factors such as interference from translations, genre or MS dating is less apparent. Still, it seems that the influence of genres is, again, weaker than the pressure exerted by the particular subperiod, as could be seen earlier from the tables 6.1.2.4e and f and as the results presented in table 6.1.2.4.g below indicate. Four samples created for this comparison pair alternately, at one time with respect to subperiodisation, at the other, the presence or absence of a uniform PTC.

Table 6.1.2.4g: Genres vs. subperiodisation, units generated on the East Midland data, sequences excluding empty categories: common base of 25 clauses

SUBPERIOD	FILENAME	GENRE	PTC[139]	F/25	AVERAGE FREQ.	CV
1350-1420	cmctmeli.m3 cmboeth.m3 cmvices4.m34 cmjulnor.m34 cmmandev.m3 cmotest.m3	philosophy/ fiction philosophy/ fiction religious treat. religious treat. travelogue bible	IS IS IR IR NN XX	5.74 4.13 6.48 6.01 5.90 5.53	5.63	14%
1350-1420	cmhilton.m34 cmcloud.m3 cmctpars.m3 cmjulnor.m34 cmvices4.m34 cmwycser.m3	religious treat. religious treat. religious treat. religious treat. religious treat. sermon	IR IR IR IR IR IR	4.29 4.08 5.59 6.01 6.48 6.06	5.42	18%
1420-1500	cmkempe.m4 cmaelr4.m4 cmfitzja.m4 cminnoce.m4 cmreynes.m4 cmcapser.m4	religious treat. religious treat. sermon sermon handbook/ other sermon	IR IR IR IR IS IR	6.09 4.93 3.77 4.35 6.25 7.98	5.56	27%
1420-1500	cmreynes.m4 cmaelr4.m4 cmcapchr.m4 cmfitzja.m4 cmreynar.m4 cmedmund. m4	handbook/ other religious treat. history sermon fiction bio/life of saints	IS IR NN IR NI NN	6.25 4.93 5.91 3.77 5.09 7.69	5.61	24%

All comparison options $p < 0.05$

The data in table 6.1.2.4.g show that even though the differences between the average frequency values are not very revealing, the level of variation (CV) does align more within subperiods (1350-1420 and 1420-1500) than within particular PTCs. In addition, the level of variation within the prototypical text category of religious instruction (IR) seems to be affected by the presence of translations (Table 6.1.2.4h). The exclusion of translated material induces a decrease in variation, typically recorded when texts classified as translations from foreign originals are taken away from samples.

‘

[139] NI: imaginative narration; IS: secular instruction; XX: unclassified PTC.

Table 6.1.2.4h: Genre groups (IR) vs. impact from translations, units generated on the East Midland data, sequences excluding empty categories: common base of 25 clauses

	SUBPERIOD	NF	AF	CV
Translated	1150-1250 1350-1420 1350-1420 1420-1500	3.31 5.59 6.48 4.93	5.08	26
Not translated	1350-1420 1350-1420 1350-1420 1420-1500	4.29 4.08 6.01 6.09	5.12	21
Both combined	1150-1500	n/a	5.10	22

Translated vs. not translated $p < 0.0001$

Unfortunately, due to insufficient amounts of data, it was not possible to verify if a similar drop in variation could be detected in the other genres. The decrease in variation which is characteristic of interference from translations, however, occurs quite frequently within various output configurations. This decrease in CV values appears to be restricted to particular dialectal areas (see also section 6.1.2.5 next).

On the whole, the difficulty in accurately assessing the impact of genres as against other factors outlined in the present section appears to spring not only from the lack of appropriate data. The attested fluidity of genres during ME might also be a problem, as remarked by Hiatt (2007). During this period, many authors tended to shift between genres or mix them (Hiatt 2007: 278, cf. Millward 1996: 188).

6.1.2.5 (Non-)pronominal SVO

A separate set of tables was generated for the SVO sequences with and without pronouns within the subject and object environments. The differentiation between inclusion and exclusion of pronominal elements from the SVO string has been crucial in the present study. At the earlier stages of development of English, pronouns would "participate in special word order patterns" (Fischer and van der Wurff 2006: 183). However, their character changed over time. The particular positional disposition of pronouns (as opposed to nominal constituents) was lost just as the subject and object function became more uniformly defined. Observing the novel behaviour of pronoun subjects and objects within SVO strings during ME, i.e. their increasing alignment with the distribution of their nominal counterparts, has been used to distinguish the emerging syntactic tendencies from the old, established structural conditions. The demise of the V2 rule from the end of ME onwards was concomitant with the obliteration of the remaining syntactic dependencies pronouns observed during the earliest days of the English language.

While in Old English a distinct positional preference was recognised with re-

spect to objects, whether pronominal or nominal, the pronominal vs. nominal characteristics of objects became less important during the Middle English period, as described in sections 4.1 through 4.3. Thus, earlier, pronominal objects tended to precede the verb whereas the nominal ones would follow the verb. Later, with the changing morphosyntactic tendencies, the positioning of the object became increasingly fixed post-verbally, thus aligning itself with the placement of nominal constituents. By comparing in this study the differences in the distribution of SVO between exclusively pronominal and nominal options for objects, it was possible to distinguish which dialect sector was the most innovative, i.e. where the postverbal position of the pronominal object would eventually be observed. With the subject constituent, too, the loss of clitic status on the subject pronouns also made nominal and pronominal elements behave alike. The two merged into a single unit. Not incidental to the entire development was the fact that the lack of clarity caused by the loss of cases on the NPs coalesced with the increasing default position of the left-most (preverbal) NP in a clause being identified as a subject (Traugott 1972: 130, Hock 1986: 369). Eventually, the pressures of generalisation (incl. subjectivisation) worked their way through to complete the syntactic change.

The opportunity to differentiate between pronominal and non-pronominal constituents of NPs also created a possibility of examining the new emerging SVO exclusively within the nominal setting as well. This, in turn, enabled observing which of the dialect sets particularly favoured the intervening verb elements between nominal subjects and nominal objects, once the overt case marking was no longer present. Establishing the limits of the link between the loss of inflections and the emergence of the rigid SVO order has been one of the aims of the present study. Indeed, the loss of case inflections on NPs has been listed among potential reasons for the turn towards the fixed VO, as presented in section 4.3.1.2. In this respect, more attention has been paid, again, to the object NPs and to the subordinate clause level, as the ones providing input on the syntactic change in question.

To gauge the extent of contact-induced pressures on the morphosyntactic change presently discussed, the issue raised in another research question of this study, the distribution of the SVO order for the non-pronominal and the pronominal option respectively, has been monitored with respect to the contrast between the Scandinavian influenced regions (the North and the East Midlands) and the rest of the dialect sectors. For the non-pronominal option, the impact of the contact with the early Scandinavian population can be found in the fast-paced inflectional erosion, especially the loss of case inflections (e.g. Iglesias-Rabade 2003). As regards the extent of the (new) pronominal layout in the emergence of SVO, the influence of Old Norse (ON) can been seen in potentially inducing particular treatment of pronominal constituents in a clause. In ON the pronouns were on the same level as full NPs. They were not clitics. Therefore, any distributional inconsistency, as recorded in the North and East Midland vs. the rest of the dialect sectors, could point to the working of this particular external pressure.

When pronouns are taken away from the SVO retrieved data pool, collectively from subject and object NP environments, only configurations with nominal constituents remain, as illustrated by the following retrieved examples:

(14) (…) and þese men preise God nyte & day (…) [ID CMCAPSER,145.20]
(15) (…) and þe hote low of þe Holy Gost openyth hyr hert, (…) [ID CMMIRK,107.2930]
(16) (…) god almichti dede werkmen into his winyarde; (…) [ID CMKENTSE,222.213]

The individual output values create wide variation in the sampling units; consider table 6.1.2.5a below. The values of the coefficient indicate a substantial diversity in the distribution of non-pronominal SVO among particular texts especially within two out of the five dialect sectors: the North and the East Midlands. Intriguingly enough, these same sectors reveal a higher or the highest overall preference for SVO (e.g. table 6.1.2a).

Table 6.1.2.5a: Comparison of the level of variation in the distribution of SVO between sequences which include and exclude the pronominal elements

CV (%)	IP- 25 IP	IP- 25 IP noPRO	IP-MAT 25 IP	IP-MAT 25 IP noPRO	IP-SUB 25 IP	IP-SUB 25 IP noPRO	IP- 200 IP	IP- 200 IP noPRO
North	6	45	26	75	26	54	6	38
East Midlands	35	51	38	48	25	62	35	51
West Midlands	34	18	14	29	53	56	34	18
Kent	15	6	9	15	31	26	15	6
South	8	31	16	22	10	41	8	31

$p < 0.0001$ for all comparisons

The decreased regularity, i.e. a high(er) CV, in these sets provides a clear signal that the increase in the variation could be a result of subtracting one of the already vital constituents of the subject and object environments from the total of all valid element possibilities within the new syntactic framework. This framework would make the subject overt and preverbal, with the object placed right after the verb. The pronominal/nominal positional distinction could no longer be as relevant as before. Furthermore, although the rise in variation occurs throughout at every clause level investigated, the CV values turn particularly high within the subordinate clause environment. It is a direct consequence of the exclusion of pronouns which at the subordinate clause level constitute a majority for the 'new' SVO in the analysed data. The tables in 6.1.2.5b-e provide a complementary portion of the output on that account.

The pronominal perspective is outlined first with respect to both subjects and objects (proSVO[140]) followed by the sequences which include either pronominal subject (proS) or object (proO). The output is available at the sentence (IP-), matrix (IP-MAT) and subordinate (IP-SUB) clause levels, with frequencies normalised to two common base options, 25 and 200 clauses.

[140] Particular sequencing options with either pronominal or nominal subjects and objects (or both) are exemplified in Appendix 3A.

Tables 6.1.2.5b-e: Pronominal SVO, sequences excluding empty categories: common bases of 25 and 200 clauses

B: Frequency

	proSVO IP25	proSVO IP200	proSVO no trans	proS IP25	proS IP200	proS no trans	proS IP-MAT IP25	proS IP-MAT IP200	proS IP-MAT no trans	proS IP-SUB IP25	proS IP-SUB IP200	proS IP-SUB no trans	proO IP25	proO IP200	proO no trans
N	1.16	8.08	8.40	3.38	25.46	28.12	3.73	28.57	27.36	3.17	25.40	28.68	1.71	13.15	14.99
EM	0.86	6.92	5.77	3.02	24.15	19.37	3.54	28.30	20.79	2.41	19.25	19.22	1.29	10.32	7.45
WM	0.86	6.99	7.13	2.72	21.72	21.61	3.25	26.02	23.56	2.32	18.56	19.57	1.42	11.35	12.28
K	0.19	1.51	n/a	2.06	16.51	n/a	2.68	21.46	n/a	1.65	13.18	n/a	0.30	2.36	n/a
S	0.54	4.32	3.93	2.77	22.14	21.26	3.58	25.17	22.01	2.15	17.19	15.85	0.98	7.86	6.82

$p < 0.0001$ for all comparison options

C: CV

	proSVO IP(25)	proSVO IP200	proSVO no trans	proS IP(25)	proS IP200	proS no trans	proS IP-MAT IP(25)	proS IP-MAT IP200	proS IP-MAT no trans	proS IP-SUB IP(25)	proS IP-SUB IP200	proS IP-SUB no trans	proO IP(25)	proO IP200	proO no trans
N	40	31	(40) 23	26	26	(14) 6	42	13	(40)17	26	29	(8) <1	23	24	(17)23
EM	54	54	38	41	41	20	45	45	35	25	25	28	62	62	30
WM	55	55	54	39	39	47	36	36	25	30	30	25	38	38	32
K	64	64	n/a	32	32	n/a	31	31	n/a	70	70	n/a	42	42	n/a
S	60	60	63	17	17	19	35	28	14	20	20	19	38	38	16

D: Frequency contd.

	proO IP-MAT IP25	proO IP-MAT IP200	proO IP-MAT no trans	proO IP-SUB IP25	proO IP-SUB IP200	proO IP-SUB no trans
N	2.23	16.01	18.44	1.39	11.36	12.69
EM	1.53	12.26	8.58	0.90	7.22	5.85
WM	1.79	14.29	13.49	1.10	8.79	10.63
K	0.44	2.05	n/a	0.17	1.35	n/a
S	1.28	9.11	7.42	0.72	5.78	4.30

E: CV contd.

	proO IP-MAT IP(25)	proO IP-MAT IP200	proO IP-MAT no trans	proO IP-SUB IP(25)	proO IP-SUB IP200	proO IP-SUB no trans
N	31	23	(29)	21	23	(17) 14
EM	67	67	39	58	58	35
WM	36	36	30	50	50	30
K	73	23	n/a	17	17	n/a
S	41	43	2	54	54	5

In addition, the configuration excluding the translated material has been added to investigate potential fluctuation of CV values resulting from the presence or absence of foreign interference ('no trans' option).

The results presented in tables 6.1.2.5b and d point to a geographically determined consistency with respect to the preference for pronouns within the new SVO layout. Whether dealing with the full pronominal SVO sequence or the element string which contains either a pronominal subject or object, the highest average normed frequencies are usually found in the North or the Midland sector. The more southern the dataset, the less pronouns are found in the new configuration. Furthermore, the shift to a higher common base (200IP) does not alter the prevalent pattern. The only change in the regional preference is recorded when the translations are excluded from the samples. In that case, the second highest average normed frequency is found not in the East but in the West Midland set. The input provided especially by the data from table 6.1.2.5b suggests that the impact of early Scandinavian structural tendencies with respect to the different treatment of pronouns could indeed be a relevant factor. Since the distribution in question is not limited to any particular clausal environment but noted throughout, the scenario of external influence looks even more feasible. With the contribution provided by language-internal pressures, one would expect a sharper contrast, for instance, between the matrix and subordinate clauses. Admittedly, the slight shift of frequency values from the prevalent pattern for the SVO order with pronominal objects (tables 6.1.2.5b and d) is striking. The difference between the East and West Midland sets is more noticeable when the samples do not contain translations from foreign originals[141].

As regards the level of variation, the values of CV indicate an overall increase in the irregularity of the distribution of pronominal SVO (Tables 6.1.2.5c and e). Much like in the reverse case presented earlier in table 6.1.2.5a, the layout of the SVO sequences are more regular when both nominal and pronominal elements are included in the subject and object categories. Further, the expected drop in variation is revealed upon the exclusion of translations. The decrease in variation appears to be more prevalent in the sector covering the SVO strings with pronominal objects rather than subjects. Along with the change in the average normed frequency values, the presence or absence of translated texts in the samples seems to create a dissonance, especially for configurations with pronominal objects.

The object unit within the SVO sequence reveals an equally interesting pattern when the focus is placed on nominal constituents (Tables 6.1.2.5f and g next page)[142]. Although the highest (average) normed frequencies are found more in the south-

---

[141] Perhaps this discrepancy is connected to the distinct assortment of genres in the samples. The West Midland set contains mostly texts belonging to the prototypical text category (PTC) of Religious Instruction (IR). The East Midland set is more heterogeneous on that account. One half of the set represents the IR PTC, the other includes, among others, the Peterborough Chronicle, viz. Non-imaginative Narration (NN), with the wild card provided by the Bible, which is mostly narrative but which also contains instructive passages.

[142] The tables 6.1.2.5f and g show the distribution of SVO with nominal subjects and objects respectively, with the sentence, matrix and subordinate clause environments distinguished. Similar to the pronominal perspective presented earlier, the tables include the frequency and CV values after the extraction of translations.

ern and West Midland dialects rather than in the sectors of the North and the East Midlands, the tendency seems to shift more in favour of the latter for the SVO sequence with nominal objects. The exclusion of translations from foreign originals appears to be the relevant factor, which introduces the shift in the regional preference for this particular configuration. In fact, when frequencies are normed to the common base of two hundred clauses (200IP) at the subordinate clause level, it is the northernmost sectors which show the highest preference for SVO with nominal objects.

Tables 6.1.2.5f -g: Non-pronominal SVO, sequences excluding empty categories, with layout after subtraction of translations: common bases of 25 and 200 clauses

F: Frequency

F/25IP	nonPROs IP-	nonPROs IP- No trans	nonPROs IP-MAT	nonPROs IP-MAT No trans	nonPROs IP-SUB	nonPROs IP-SUB No trans	nonPROo IP-	nonPROo IP- No trans	nonPROo IP-MAT	nonPROo IP-MAT No trans	nonPROo IP-SUB	nonPROo IP-SUB No trans	nonPROo IP-SUB 200IP	nonPROo IP-SUB No trans
N	2.07	1.65	4.26	3.03	0.96	0.83	3.73	3.50	5.76	4.75	2.71	2.77	23.03	25.72
EM	2.15	2.07	3.18	2.84	1.20	1.35	3.84	3.62	5.06	4.44	2.66	2.98	21.28	23.86
WM	2.12	2.30	3.44	3.35	1.10	1.40	3.44	3.47	4.92	4.56	2.30	2.52	18.42	20.19
K	2.45	n/a	4.27	n/a	1.24	n/a	4.32	n/a	6.47	n/a	2.88	n/a	23.04	n/a
S	2.76	2.85	3.95	4.16	1.46	1.54	4.55	4.65	6.24	6.53	2.90	2.94	23.16	23.53

$p < 0.0001$ for all comparisons

G: CV

CV	nonPROs IP-	nonPROs IP- No trans	nonPROs IP-MAT	nonPROs IP-MAT No trans	nonPROs IP-SUB	nonPROs IP-SUB No trans	nonPROo IP-	nonPROo IP- No trans	nonPROo IP-MAT	nonPROo IP-MAT No trans	nonPROo IP-SUB	nonPROo IP-SUB No trans	nonPROo IP-SUB 200IP	nonPROo IP-SUB No trans
N	40	41	67	24	59	105	10	10	39	31	29	31	26	15
EM	47	64	46	65	56	51	32	41	36	58	24	21	24	21
WM	16	7	21	31	44	30	17	21	14	11	33	44	33	37
K	6	n/a	23	n/a	26	n/a	18	n/a	16	n/a	30	n/a	30	n/a
S	28	36	20	26	36	64	13	15	12	19	15	15	15	21

Yet another important observation can be made concerning the level of variation in the distribution of non-pronominal SVO (Table 6.1.2.5g). In this case, too, a characteristic layout is found in the nominal object more often than in the nominal subject section. When analysing the CV values at the sentence level, the SVO distribution with nominal objects turns out to be distinctly regular (figures shaded in green). Four out of the five dialect sets in PPCME2 exhibit the level of variation below 20, which is actually one of the lowest (i.e. most regular) overall scores recorded within all configurations in the present study. The distribution of SVO with nominal objects is more regular than that of the VO pair (see next section) as well as the SVO string which incorporates both nominal and pronominal elements. The peculiar regularity of SVO with nominal objects seems to conform to the distributional patterns dating to the earliest days of English when nominal objects tended to be mostly postverbal (e.g. Mitchell and Robinson 1983: 61, Fischer *et al.* 2000: 161). It appears that later on the development became established and, as the results from the previous set of tables show, extended to the distribution of pronouns. Further, a frequently recorded drop in variation upon the exclusion of translations is present here as well. The decrease in CV values is almost exclusively limited to the North and the East Midland pair, a tendency already revealed within the distribution of SVO with both nominal and pronominal elements (tables in section 6.1.2.3). Should the growing preference for nominal objects to follow the verb be associated with the loss of cases on NPs, the decreasing CV values as found in the northernmost sectors appear to be quite in place. These sectors represent dialects that were the first to shed the inflectional endings. Lastly, the fact that the pattern for SVO with nominal objects shifts specifically at the subordinate clause level and not the matrix one also seems quite pertinent. The loss of the preverbal position of objects in ME has been shown to be directly associated with the loss of 'subordinate', viz. verb-final, word order (e.g. Traugott 1972: 161).

6.2 DRAWING A WAVE PATTERN

As outlined in 5.1, one of the goals of the present study has been to establish to what extent morphological simplification, more specifically the loss of cases on NPs in early English, had conditioned the concomitant syntactic shift to the rigid SVO word order. The most obvious reason for linking the two developments is the observation that in the early stages of development of languages case systems helped to differentiate subjects from objects (e.g. Robinson 1992: 166). Indeed, the earliest examples of SVO in OE are those which alleviated potential ambiguity. This particular word order sequence was valid especially with caseless nouns in the function of subjects and objects (e.g. Bean 1983: 139). In addition, sections 3.1 and 3.2 showed that there is, indeed, a correlation between word order freedom and presence of an overt case system. Once the case markers are not available, the other means of identifying grammatical categories usually needs to be employed, i.e. verb agreement and stricter word order. English is known to stand out in the Germanic language family as the one using rigid SVO as a primary means to distinguish basic grammatical relations in a clause. As regards early English in particular, section 4.3 highlighted instances in which morphological sim-

plification, case loss included, contributed to the arrival of new syntactic tendencies. They encompassed the emergence of overt SV, the loss of clitic status on pronouns and the shift of the object from preverbal (OV) to postverbal (VO) position.

The second important aspect explored in the current research has been to estimate the role of the Anglo-Scandinavian contact in the morphosyntactic change here discussed. The erosion of (case) inflections has often been attributed to this particular contact situation (e.g. Iglesias-Rábade 2003: 86). In addition, some other NP oriented developments, notably the loss of clitic status on pronouns, which had an effect on both subjects and objects as well as on the environment around VP, have been identified as Scandinavian contact-induced phenomena (e.g. Fischer and van der Wurff 2006: 185). The contact scenario currently proposed has been that of convergence, whereby the (phonological, lexical as well as grammatical) inventories of the languages in contact become increasingly similar (e.g. Matras 2009: 230). On that scenario, the erosion of inflections in early English, most prominent in the areas affected by contact, could be viewed as the strategy to eliminate contrasts between the two (already similar) linguistic systems. On the dialectal map of Middle English the effects of convergence would have to be manifested by the increased preference for (S)VO especially in the sectors which correspond to the regions previously held under Scandinavian control (i.e. the Danelaw), viz. the North and the East Midlands (e.g. Townend 2006: 65). Indeed, as remarked by Corrie (2006), contact with the early Scandinavian language was "with a few important exceptions, a geographically specific phenomenon". This fact clearly differentiated contact with Norse from other external influences on English at that time, notably Latin (Corrie 2006: 97).

Exploring the distinction between the innovative sectors (those where SVO tokens were most numerous) and the other dialect sets (where feature was less pronounced) has been found in the working of the convergence model, the wave theory. The theory was chosen to facilitate the explanation of the changes presently discussed. The model involves the idea that there is an influential centre from which the new feature radiates towards more outlying areas. This centre is commonly referred to as the focal area, a location which usually designates a cultural centre of some sort or an area of political dominance (McMahon 1994: 229). In the current study, the nexus between the two processes, the inflectional erosion and subsequent emergence of SVO is located in the areas of intense language contact between English and Scandinavian (earlier Celtic) populations. From these locations, the 'new' feature diffused through the socio-geographic space, over a longer period of time largely by population mobility. The following sections treat on how well we could distinguish these aspects of the wave theory.

6.2.1 The Danelaw and the focal area

When analysing the results of this study, it has been repeatedly recorded that the new syntactic tendencies tend to prevail most often in two dialects of Middle English – the North and the East Midlands. Firstly, both dialects exhibit the most prevalent preference for SVO at the IP-SUB level. The highest average (normed) frequencies are maintained in these sectors regardless of the sequencing option and the common base adopted (see Table 6.1.2.1b). The popularity of SVO within the subordinate

clause environment is even strengthened if translations are excluded (Table 6.1.2.3a). Furthermore, when examining the impact of foreign translations alone, a frequently occurring drop in variation has been recorded once the attested translations have been taken away from the samples. Intriguingly enough, the decrease in variation usually took place in these two northernmost dialects. Further, the comparison of the earlier and later portions of the ME material suggests that SVO was more established in the North and the East Midlands already during the first half of the period (tables 6.1.2.2a and b). Secondly, as regards the nominal and pronominal distinction, the older state of affairs, namely a frequent SVO order with non-pronominal objects, seems to be an established phenomenon, as could be expected. All the dialects show relatively low CV values. The most regular distribution, however, is displayed in the North (table 6.1.2.5g). Thirdly, after the extraction of translated material, the highest frequency values for SVO at IP-SUB are recorded in the East Midlands. When the common base is changed to 200IP, both the North and the East Midlands exhibit the highest preference for SVO when the object is not a pronoun. A typical decrease in variation upon the removal of translations has been distinguished here as well. Fourthly, for the SVO sequences with pronominal objects, the highest average frequencies have been recorded in the northern set throughout (overall, matrix and subordinate clause environments respectively), followed immediately by the Midlands (the West more than the East). It seems that the postverbal position was extended from nominal to pronominal objects. Altogether, the more north one heads on the dialectal map of ME (table 6.1.2.5b and d), the more differently pronouns tend to behave (i.e. resemble full NPs). Finally, the assessment of the level of variation across the board has shown that the two sectors, the North and the East Midlands, more often than not pair up in their values, as if particular tendencies affected them collectively[143].

Explanations for this overt preference for new structural tendencies manifested in the North and the East Midlands can be many. Yet, it appears that this pattern would least likely be explained by general, ongoing tendencies in the language, as regulated solely by the internal pressures. Too many output configurations point to a specific location on the dialectal map of ME for that to be a mere coincidence. At this point, the reasons for the change need to be sought within the working of external pressures, such as language contacts. The Scandinavian input on that account seems to be the most fitting. The correspondence of the two dialects with the areas of the early Anglo-Scandinavian contact is evident. The longstanding contact with the early Scandinavian population, as illustrated in Table 6.2.1a below, can indeed be considered a likely candidate for the catalyst, the pressure reinforcing this kind of change in morphosyntax:

[143] E.g. CV for SVO within IP-SUB environment, especially for higher common bases and after the extraction of translations the values of both drop to the same level; CV for SVO with non-pronominal objects within IP-MAT as well as IP-SUB; CV for VO within IP-MAT environment.

Table 6.2.1a: Comparison of syntactic tendencies in OE and ME, with developments characteristic of Anglo-Scandinavian contact zones emphasised

NEW SYNTACTIC TENDENCIES IN MIDDLE ENGLISH	DEVELOPMENTS PREVALENT WITHIN SCANDINAVIAN CONTACT ZONES	SYNTACTIC TENDENCIES IN OLD NORSE
• SVO in both IP-MAT (+ V2) and IP-SUB (no V2) • Postverbal position assumed for both nominal and pronominal objects • Pronouns no longer clitics (positionally integrated with full NPs) • Almost complete loss of inflectional morphology	• SVO in both IP-MAT (+ V2) and IP-SUB (preference for the latter clearly geographically defined) • Highest preference for both nominal and pronominal objects placed postverbally • Pronouns no longer clitics – positionally more integrated with full NPs • Loss of inflectional morphology first attested and completed	• V2 language with SVO in IP-MAT (always), in IP-SUB (normally)[144] • Pronouns behave like full NPs (not clitics)

To add to the list of phenomena involved in typical contact scenarios such as morphological simplification and the turn towards SVO, the influence from the early Scandinavian language introduced the availability of properties characteristic of ON. The particular treatment of pronouns, as full NPs, too, seems to have pushed the transformation of structures in early English into a new direction. On the whole, the observation made during the preliminary stages of the present research, namely, that the North and the East Midlands "form a truly Scandinavian contingent" (Czerniak 2011: 149) still stands. Only the configurations to retrieve output from specific clause levels were investigated at that point. Subsequent analysis of the data, which explored the problems around MS dating as well as the interference from translations and distinct pronominal distribution, as the above outline reveals, clearly substantiates that claim[145].

Interestingly, while in many configurations the new syntactic layout is clearly favoured in the entire contact area, viz. the North and the East Midlands, it is the North where preference for the investigated feature is the most prominent and most regularly distributed. It would follow, therefore, that within the Anglo-Scandinavian contact zones there is a division into areas more and less 'disposed' to these external linguistic pressures. Using the wave model nomenclature, the location of the influential centre (i.e. focal area) extends only to a specific section of the region which executed Scandinavian laws and customs. By legislation and prevailing habits, the Danelaw could be regarded as a single unit. Linguistically, however, there seems to be a difference in volume of how features of Scandinavian provenance integrated into the early English language, as noted, for example, by Samuels (1985). The difference in question corresponds to the borderline situated on the river Humber (section 2.3).

---

[144] As quoted from Swan (1991: 234-6).

[145] The pilot study (Czerniak 2011: 139-154) included analysis of SOV-SVO, OV-VO variables. In the end, the final cuts on the samples involved only SVO and VO sequences, with the additional factors explored. Moreover, the error concerning the scope of VP has been limited to the minimum (see Appendix 5A).

Samuels (1985) recognises this difference and claims that the Dano-Norwegian settlements north of the river "must have been of a kind that was denser and [that] brought about a deeper linguistic penetration" than the Danish settlements on the southern side of the river (Samuels 1985: 271-2).

Although the place-name evidence would not directly support that claim, there are a few reasons in favour of the Northern-Southern Danelaw partition. Above all, there is the issue of spoken Scandinavian language surviving longer north of the Humber (Samuels 1985: 272, cf. Bugge 1921: 182-6). The North proved to be a peculiar (i.e. difficult?) case in other respects as well. Just as ON enjoyed a persistent following so were the Scandinavian customs, which lingered there for a time (Roesdahl 1991: 248). "A deep-rooted political separatism" of the North, too, must have fuelled the strength of the Scandinavian impact (Higham 1993: 198). Yet, the unwavering presence of the Scandinavian language and traditions in the deep North did not exclude the possibility of distinctive interpenetration of English and Norse in these areas. Barnes (1993) provides a scenario for such a state of affairs:

> *A desire arose among the Scandinavians to distance themselves from their sordid past and to adapt to the higher and more stable Anglo-Saxon culture. The product of this attitudinal shift was a process of Anglo-Scandinavian acculturation, thoughtful and deliberate, one of whose manifestations, if not the most visible, was the establishment of a higher-level form of Anglo-Scandinavian speech, which served to make the settlers feel more at home in their new surroundings while still emphasising their separate identity (Barnes 1993: 71).*

In his attempt to identify the focal area Samuels emphasises the importance of Lincolnshire (Samuels 1985: 275). This region could be included in the areas of deeper linguistic (Scandinavian) penetration. Other scholars, too, confirm the existence of such a centre of influence. Geipel (1971) and Millward (1996) are just some of the many who subscribe to the idea that the North was a much more progressive region linguistically speaking than the rest of the country (Geipel 1971: 58-9, Millward 1996: 142).

In the output charts (section 6.1.2), the preference for SVO was always most prominent and most regular in the North than the rest of the sectors. Admittedly, the East Midland set, the other part of the Scandinavian contact zone, often aligned itself with the North in the higher or highest average (normed) frequency values, especially at the subordinate clause level (table 6.1.2.1b). This very tendency lingered after the exclusion of the foreign translations, which only strengthened the validity of the Scandinavian impact (table 6.1.2.3a). Still, having compared all the available output configurations, it is the northern set which exhibited the preference for the new syntactic framework most clearly. The lowest level of variation displayed in this dialect was exceptional. The regularity of the distribution of the feature continuously maintained at the level of 10 per cent or below (e.g. tables 6.1.2b; 6.1.2.1c; 6.1.2.2c; 6.1.2.3b and c; 6.1.2.4d; 6.1.2.5g – for nominal objects) points to the uniformity between the texts of the set which surpasses the impact of any other influencing factor investigated in the current study. Such regularity suggests that the popularity of SVO as displayed within one text of the northern dialect falls almost identically in line with the pattern

of the distribution of this feature in other texts in the set. Even if the aforementioned underrepresentation of texts time-wise was taken into account, the North would still be the most innovative dialect within the entire PPCME2 material. As set against the southern sector, which also contained data predominantly from the second part of ME, the North continued to show the highest regularity (i.e. the lowest CV values) of the feature for both sequencing options, viz. SVO strings with and without the empty elements (e.g. table 6.1.2b). This pattern clearly distinguishes the northern dialect from the South or any other sector of the database for that matter. Considering the North as the likely candidate for the centre of the linguistic influence during ME, the patent prevalence of SVO converges not only with the noticeable distribution of pronouns in the new configuration but also with the commonly accepted onset of inflectional erosion.

An attempt at estimating the role of morphological simplification in the emergence of the new syntactic framework in early English has been achieved, also, by examining the general distribution of SVO at different clause levels and by subsequently comparing it to the results of the distribution of the VO constituent pair. Tables 6.2.1b and c below outline the distribution of SVO and VO respectively within particular clause environments, with frequencies normed to the common base of 25 IP. The absolute frequencies have been added to both tables to ensure an accurate proportional relation between the instances of VO and SVO[146].

As indicated by the figures in both tables, the layout of SVO largely conforms to the distribution of the VO pair. A full correspondence of the normed frequencies is especially visible overall, i.e. looking at the total of all matrix and subordinate clauses (IP-). For matrix clauses, the highest preference for (S)VO changes only between the North and the South. At the subordinate clause level, the lowest preference for the two element strings shifts solely between the West Midlands and Kent. Apart from the expected higher normed frequencies for the VO pair, there is an increase in the regularity of the feature recorded when compared with SVO. The decrease in variation is recorded in two thirds of all cases and in one other the increase is by 1 per cent from a very regular 6 per cent to an equally regular 7 per cent, which suggests that the VO pair is, on the whole, more established than SVO.

Table 6.2.1b: Distribution of SVO at all clause levels, sequences without empty categories. Data copied from the tables in sections 6.1.2 and 6.1.2.1: common base of 25 clauses

SVO	**IP-MAT**			**IP-SUB**			**IP-**		
	Raw count	F/25	CV	Raw count	F/25	CV	Raw count	F/25	CV
N	988	7.98	26	999	4.27	26	1987	5.56	6
EM	2120	6.68	38	721	3.72	25	2841	5.24	35
WM	3481	6.73	14	2179	3.25	53	5660	4.61	34
K	808	6.94	9	453	3.01	31	1261	4.62	15
S	2307	7.55	16	828	3.72	10	3135	5.60	8

[146] It is expected that the number of VO instances outnumber those of SVO.

Table 6.2.1c: Distribution of VO at all clause levels, sequences without empty categories: common bases of 25 clauses

VO	IP-MAT			IP-SUB			IP-		
	Raw count	F/25	CV	Raw count	F/25	CV	Raw count	F/25	CV
N	1166	9.25	21	1541	6.75	17	2707	7.61	7
EM	2528	8.32	25	1422	6.52	28	3950	7.39	26
WM	4796	8.72	7	3670	5.58	41	8466	6.77	30
K	931	8.46	15	1001	5.86	14	1932	6.93	14
S	3274	9.74	14	1494	6.36	22	4768	7.98	18

SVO vs. VO overall $p < 0.0027$
other comparisons $p < 0.0001$

The high(er) preference for VO at both matrix and subordinate clause levels indicates that the shift to the postverbal position of objects took place during the Middle English period.

When looking at the differences in raw count proportions of both SVO and VO dialect by dialect, it becomes clear that the most likely element appearing on the left side of the VO is, indeed, the subject. The proportion of SVO in the VO pool is expectedly greater for the North than for the South, as shown in table 6.2.1d below. What is more, the two dialects differ most in their proportions of SVO at the subordinate clause level, which is not surprising either.

Table 6.2.1d: Proportion of SVO in VO for North and South - Calculations made on the absolute frequencies

% on raw count	NORTH (N)	Difference between N & S	SOUTH (S)
IP-	73 %	7 %	66 %
IP-MAT	85 %	1 %	84 %
IP-SUB	65 %	10%	55 %

The more conservative South seems to have accepted the SVO order within this clausal environment at a slower pace than the innovating North. The highest proportion of SVO at the matrix clause level in general can also be easily accounted for. These clauses exhibited an SVO grammar both with and without V2 (details in section 3.1). The emerging structural tendencies only added more instances of the feature, although the rules of creating SVO at that point would obviously be different. Further, analysing differences in the values of the coefficient of variation for these two geographically distinct sectors reveals a particularly close alignment of SVO and VO distributions in the North overall (Table 6.2.1e below). This overt regularity of SVO and VO, with the fluctuations reaching only 6 per cent and 7 per cent[147], respectively, suggests that the dispersal of the two configurations could have been regulated by

[147] Hence 1 per cent difference.

the same mechanism. Furthermore, the level of variation recorded for the North contrasts, in fact, not only with the CV values recorded in the South but also with those recorded in the remaining dialect sets.

Table 6.2.1e: Differences in CV between SVO and VO for North and South – CV values expressed in percentages by default

	NORTH	SOUTH
IP-	1 %	10%
IP-MAT	4 %	2 %
IP-SUB	9 %	12 %

The North seems to be very much fixed on the tendency to have postverbal objects just as it is focused on the placement of the subjects preverbally. The configuration employed already during OE times to alleviate potential ambiguity appears to be substantially established in the very same dialect which first shed the inflectional endings. That we are dealing with a new structural development is demonstrated by the overt tendency for the feature to appear at the subordinate clause level. The high average normed frequencies in the North for this clausal environment are maintained with low and high common bases for both sequencing options (with and without the empty categories). The exclusion of translations only strengthens the pattern.

6.2.2 Pace of the spread towards the end of Middle English

Localising the focal area provides only a part of the wave theory. The model also needs to account for the so-called contagious diffusion of the novel feature off the centre of innovation to the surrounding areas, to the peripheries, as outlined in section 3.4. Indeed, scholars investigating the Anglo-Scandinavian contacts have mentioned the second phase of impact, the one which involved the spread of the novel features via interdialectal contact (e.g. Hogg and Denison 2006b: 15). It is through this means that the change originating at one locale could percolate to more outlying linguistic regions, eventually hitting the South. Although the wave framework can be applied to the study of grammatical features, it has been largely limited to the study of phonological variables. Using it to explain the route of distribution of overarching grammatical patterns might seem to be quite challenging. One could wonder whether it is at all possible to identify a more or less uniform path of spread of something so complex as word order, especially considering the period during which the changes began to transpire. After all, nothing firm can be stated about Middle English other than the fact that it was the time of transition between old and new in terms of linguistic features and preferences. Not even clear dialectal boundaries can be confidently drawn (e.g. see Wales 2006: 34-6). SVO was only beginning to dominate over other word order possibilities. It took a few more centuries before it became unquestionably dominant. Nevertheless, particular text specimens from the period, the works of Chaucer for example, can provide us with useful hints with respect to the progress with which the changes proceeded. In her analysis of Chaucer's late 14th century material Corrie (2006) reveals that both new morphosyntactic tendencies as well as occasional returns

to the older state of affairs are present in his texts. Thus, there are prepositional phrases taking the place of OE inflections, with the word order leaning more clearly towards SVO. On the other hand, the Old English SOV order can be detected as well, along with the conscious use of *–e* inflections of the OE weak/strong adjective system (Corrie 2006: 107-8, also Horobin and Smith 2002: 138). It is common knowledge that Chaucer's English, i.e. the language of London, gave the primary basis for the emerging standard, modelled on the EM dialect.

Intriguingly enough, examining London English can also tell much about both the route of the spread and potential transmitters of the features which were soon to become prevalent. Firstly, the variety in question, apart from its clear East Midland provenance, also included some of the typically northern elements. These features, as remarked by Knowles (1997), would spread from the Danelaw area (Knowles 1997: 45, 53). Secondly, if we were to follow Samuels' classification (1963) of incipient standard varieties present at that time, a clear distinction could be made, among others, into an early and later London dialect. It was the latter which was influenced by population movement from the central Midlands (Knowles 1997: 53). Knowles's subsequent comments on the status of London English can, in fact, provide us with the details which set almost a prototypical wave pattern. Thus, he claims that "the southward movement of northern and east midland forms into London points to a prestigious centre further north, presumably York" (Knowles 1997: 127). Obviously, reality imposes checks on idealistic lines drawn by theoretical models. While London can ultimately be regarded as the destination point of the spread of innovations, the north-to-south path may not necessarily be applicable to most of the available novel features which were added to the morphosyntactic pool of English (cf. Tristram 2007: 207-8, cf. Samuels 1985: 269).

The wave model works but instead of one geographical focus there could be several different foci, as is suggested by, e.g. Hudson (1980: 40). Nevertheless, my study on the emerging (S)VO has shown that there is a parallel between the 'ideal' case (one focal area) and the state of (apparent) linguistic reality depicted in the Middle English corpus (PPCME2). New syntactic conditions almost always radiate from the north southwards. The exceptions to that pattern include a high frequency but not a low CV for the South overall (total of IP-MATs and IP-SUBs). As mentioned in section 6.1.2.1, the boost is likely to be generated due to the availability of southern texts exclusively from the second part of ME. Different sequencing options as well as other factors potentially influencing the distribution of SVO do not cause so much interference as to shift the route of the spread of the investigated feature. The exclusion of translated material, as presented in section 6.1.2.3, only enhances that pattern. The differentiation between the earlier and the later material (section 6.1.2.2), too, clearly points to the feature being more established during the latter part of ME, with the prevalence touching larger areas of linguistic terrain. This correlates with various studies which indicate that although SVO predominates in ME texts right from the beginning (e.g. Fischer *et al.* 2000: 206), it was after 1300 that the increase in VO is quite evident, notably at the expense of the earlier OV (e.g. Fischer *et al.* 2000: 162). Preference for a 'new' SVO is usually most evident, as indicated by high frequencies and low CV values, in the northernmost sets. From there the preference decreases southwards. The north-to-south direction of the spread, as well as the progressive nature of the northern

dialect over the others can already be identified when comparing dialect translations included in PPCME2 (section 6.1.2.4). Interestingly, the <cmedthor> and <cmedvern> pair very accurately mirror the patterns found in the entire dialect sets. It is a clear signal of how firmly established SVO actually was at that time. It also suggests that the material in the Middle English parsed corpus is, indeed, representative, despite some of its shortcomings. Furthermore, the particularly high frequency of SVO in subordinate clauses in the northern version underlines its higher/highest overall frequency. This fact places the North ahead of the rest of the dialects. It, too, shows that the growing prevalence of SVO was advancing, as the new syntactic conditions were detected at both sentence levels in that dialect.

As for the ideal case of the 'contagious diffusion' of the investigated feature, the frequencies of occurrence and the values of the coefficient of variation make each of the dialects of the corpus unique. The dialect uniqueness could manifest the advancement of change, as remarked by McMahon (1994: 228). On a linguistic map of ME, there is always a region where SVO is preferred more or less explicitly. Admittedly, the clean-cut progression of the change by stages in time through socio-geographic space, advocated by the wave model, is not neatly mirrored in the spread of the new syntactic tendencies at issue. Yet, with respect to the factors presently discussed, one can distinguish both the focal area, from which the structural innovations radiated and the peripheries, i.e. the regions which accepted these innovations at a slightly later date. The wavelike ripples, perhaps, would not be laid out in such a uniform fashion, much due to the fact that, apart from the developing new pattern handled in the current research, there were many other secondary developments. They, too, complemented the emerging of new syntactic tendencies and they, undoubtedly, had their own unique path of spread. At the very least, a sharp division between the matrix and subordinate clause levels thrived well after the Middle English period.

In all, by the end of the ME the general characteristics of Modern English word order were established. The following centuries witnessed only the strengthening of these characteristics (e.g. the rise of DO-support). The competing OV variant disappeared in prose around the mid-sixteenth century (Fischer and van der Wurff 2006: 184, 186-7). The typically Germanic V2 became marginalised some time during the seventeenth century (Fischer and van der Wurff 2006: 184).

7 Discussion and Conclusion

The current, last chapter starts by consolidating and interpreting the results presented in chapter 6. Some of the issues raised in the socio-historical part (chapter 2) will be addressed and reviewed in light of the linguistic evidence. Section 7.2 evaluates the contribution of this study to the research within the area of early English syntax. The discussion ends with considerations and possible directions for further research on the emergence of the SVO word order in English.

7.1 SUMMARY OF THE MAIN FINDINGS

The present study has dealt with one of the most important morphosyntactic changes occurring at the early stages of the development of the English language. The result of that change was the emergence of the rigid SVO order as the major determinant of grammatical relations in the sentence, the structural tendency which has thrived up to now. In this respect, English is unique within the entire Germanic family in having predominantly this syntactic strategy at its disposal to differentiate subjects from objects. As described in section 3.1, there are two other languages in the family, Dutch and Frisian, which use only one way to distinguish grammatical relations. Yet, they do it by means of verb agreement, not by a strict word order.

As English experienced a thorough morphological simplification concurrently with the word order shift and since a causal relationship between the loss of case distinctions, in particular, and the emergence of SVO has been suggested in the literature (sections 3.1-2), the current research has explored the extent of that dependency. One of the influences behind the loss of cases and the concomitant turn towards strict word order was the series of language contact situations which English went through during its early phases of evolution. The present study has aimed at exploring the nature and the importance of the contact the English people had with the Scandinavian population. This was one of the most significant external impacts *Vox Anglica* had undergone early on, affecting not only its lexico-semantic domain but also that of syntax. The contact with the early Scandinavians has been regarded as instrumental at least in the process of attrition of case inflections in English.

Since the convergence scenario, the merging of the two systems, has been adopted to explain the path of the morphosyntactic change at issue, this study has tested the applicability of a theoretical model which should facilitate the explanation of the changes tackled in the current study, viz. the wave theory. This model was chosen not only on account of it being pertinent to the contact circumstances but also because it emphasises the two aspects that are vital in tracking linguistic changes through a socio-geographic dimension, namely time and space. These two aspects were explored in the databases used in the current research – *The York-Toronto-Helsinki Parsed Corpus of Old English Prose* (YCOE) and *the Penn Parsed Corpus of Middle English,*

Second Edition (PPCME2). Apart from shedding light on the nature of the early morphosyntactic change presently discussed, the analysis of the diachronic corpora has served to validate their applicability in complex studies of such overarching features as word order.

As regards the first research question, viz., to what extent morphological simplification conditioned the syntactic change leading to prevalent SVO, the present study has aimed at confirming the existence of a convergence point on a dialectal map of the parsed corpora material between the inflectional erosion and the emerging SVO. If such a point could be identified, could both processes be considered interdependent and thus suggest that the convergence point marked the point of origin of the new syntactic tendencies? Taking into account the evidence from the earliest English texts, the convergence point would have to be confined to the early English northern dialect zone. It is the place where the inflectional decay was most advanced. Indeed, every data sample analysis in this research to retrieve the proportion of (S)VO revealed the highest and, for the most part, the most regular preference for the feature in this particular dialect sector. Firstly, the comparison of the syntactic environments between the Old and Middle English periods showed that the biggest growth of SVO structures occurred in the northernmost sets, i.e. those which represent dialects that evolved from OE Anglian into ME Northern and East Midland sets. The significance of that finding is strengthened by the fact that the increase in the preference for SVO was not markedly conditioned by pressures regulated at particular syntactic levels, matrix vs. subordinate clauses, or time frame normalisation or the presence of foreign translations, all of which have been found to have bearing on the word order distributions. Furthermore, considering the fact that the very first instances of SVO in Old English were precisely those which enabled differentiating caseless subjects from objects (e.g. Bean 1983: 139), the increase in SVO in the dialect sets which first showed substantial loss of case inflections was to be expected. Secondly, during the Middle English period alone, the North manifested itself as the dialect where the high(est) normed frequency and the most regular distribution (shown by a low CV value) of (S)VO were recorded. The differentiation into element sequences which either include or exclude empty categories did not change the output. The tendency was even strengthened when empty categories were taken away from element strings.

The other dialects tended to follow the North with a more or less regularly advancing geographic distribution, at some point even approaching the mode of what Wolfram and Schilling-Estes (1998) have referred to as "contagious diffusion" (Wolfram and Shilling-Estes 1998: 143). The East Midland set and the South showed parallel normed frequency values. Still, the rate of repetition of the highest and the most regular preferences for (S)VO within different sampling strategies never reached that of the North. At the same time, it is tempting to link the uniformly conservative character of Kent and the West Midlands, manifested by the lowest preference for (S)VO (Table 6.1.2a), directly with their rather determined resistance to morphological simplification. Both dialects were the last ones to accept the loss of differentiation between NOM and ACC singular forms on nouns and adjectives just as they kept on preserving the weak/strong, singular/plural distinctions on the latter (Fig 4.2.2a). The fact that the bastion of the early Anglo-Saxon literary tradition was established in the south-west-

ern Midland area was undoubtedly instrumental in causing this delay. Furthermore, when analysing the distribution of the investigated feature within particular clausal environments, the northern set displayed the high(est) frequency values for both matrix and subordinate clauses, although the distribution in the former was not so transparent as that of the latter. The disparity can be confidently attributed to the presence of the V2 tendencies at the matrix clause level. As mentioned in section 3.1, SVO with V2 grammar produces the same surface order as an SVO grammar without this constraint.

The subordinate clause setting, on the other hand, where SVO began to supersede the Old English SOV, revealed the distribution of the feature as a genuine innovation, free from the V2 pressures. Here, the preference for the new syntax was the highest and most regularly distributed first in the North, with the East Midland dialect following suit. The particular regional preference at the subordinate clause level not only aligned with the morphological simplification locale but also suggests a more external rather than internal motivation for the structural change. Furthermore, comparing the earlier and later Middle English data, the northernmost sectors expectedly showed the highest normed frequencies for SVO for both, with the trend on the increase over time (Tables 6.1.2.2a-b). The entrenchment of the feature was explicitly indicated throughout by the low(er) values of the coefficient of variation during the second half of the ME period (Tables 6.1.2.2c-d). Moreover, the distribution of normed frequencies for the second part of the period could already indicate two separate influencing forces: on one hand, the spreading innovations from the North, and on the other, the growing power of the South after the reestablishment of the English identity and language. The North-South difference in the distribution of SVO was also visible when it was compared with the distribution of the VO pair alone, as presented in section 6.2.1. Most of the dispersal of VO reproduced the pattern of SVO with the highest normed frequency values fluctuating between the two dialects, especially at the matrix clause level. Yet, the highest overall regularity of the distribution of both sequencing options was only maintained in the North, which shows that the postverbal position of objects, the development prompted by the loss of cases on NPs, started off in this very dialect.

In all, the corpus data analysis suggests that the particular preference for (S)VO and the morphological simplification converged in the same location. A random output configuration is hardly a factor here, since the distribution of the feature repeatedly produced the expected layout. The North was where both the loss of (case) inflections and the onset of new syntactic conditions meet, while no other dialect approached the volume of that alignment. Given that many developments which accompanied the emergence of SVO were reinforced by the loss of inflections (section 4.3.1) and that the latter could induce identification problems on NPs, the connection between the two seems quite straightforward.

The second question tackled in the present research has addressed the role of the early Anglo-Scandinavian language contacts in the morphosyntactic change under discussion. Here, the corpus study has aimed at establishing whether the convergence point mentioned earlier correlates with the dialects of the areas affected by the contact, i.e. the North and the East Midlands. It seemed at least pertinent to investigate

the emergence of SVO from that particular angle, since the loss of case inflections has often been attributed to this contact situation (section 4.4.2). Additionally, as morphological simplification, serving as a means of eliminating contrasts, is an attested language contact effect, the convergence scenario is plausible enough. Change towards SVO itself, too, has been commonly regarded as one of the frequently occurring contact-induced outcomes (section 3.3.2). As for the Scandinavian impact against other foreign influences, the contact with the Norse language was, indeed, a geographically defined occurrence. With the exception of the Celtic input, present in the North (and the West Midlands), other external influences, such as Latin and French, were more passive and stylistic in nature, affecting any dialect which had contact with foreign literature.

The substantial growth rate of the subject-verb-object order identified in the former Anglian and subsequent northern and Midland sectors points in itself to a common setting between the emergence of the new syntax and the relevant contact zones. Yet, it is the output of the Middle English period that provides an explicit indication of the contact-induced pressures. After all, we cannot confidently speak of the spreading of Scandinavian forms in writing until the late eleventh and the early twelfth century, even if these would be, in fact, a belated echo of these same developments long entrenched in speech (e.g. Horobin and Smith 2002: 129).

The results showed that apart from the overall high and very regularly distributed SVO in the North, there was a marked preference for the feature predominantly at the subordinate clause level (IP-SUB) in both the North and the East Midlands. The same dispersal was replicated, more decidedly, in the VO configuration. Furthermore, the level of variation (CV) for the SVO sequence in IP-SUB was maintained at the same level in both sectors. In fact, the tendency to show parallel levels of variation for the North and the East Midlands has very often been demonstrated in this study in various sample arrangements. The two dialects paired off quite consistently, especially at the subordinate clause level and after modifying the data sets in order to account for different influencing factors. The particular, regionally conditioned and uniform distribution of SVO at IP-SUB within both dialects is not only a clear indicator of language contact pressures with the Scandinavian population but also a signal of the development being the consequence of other linguistic pressures, not solely the modification of basic syntactic patterns based on those found in the fellow contact language (cf. Lehmann 1973: 201). Thus, instead of leaning more towards the Scandinavian syntax, a combination of the familiar V2 rule and SVO active in both clausal environments, English adopted a rigid SVO basic word order and soon restricted the applicability of the former, typically Germanic feature.

The changes which contributed to the establishment of the rigid SVO must also have been associated with the NP setting, where the roles of constituents acting as subjects and objects needed to be invariably explicit in the absence of case marking. The results showed that of the two Scandinavian-influenced dialects the northern set displayed both a high and most regular overall preference for SVO, linking the shift to SVO directly with the loss of case inflections. Admittedly, the language external contribution, at that point, cannot have been exclusively limited to the Scandinavian provenance only, especially if the comparison of the OE vs. the ME material is taken

into account. It is possible that Celtic influence, dated already to the pre-Viking times also played a role in determining the extent and pace of morphological simplification (see e.g. Kastovsky 1994). Moreover, Trudgill (2010: 1-35) discusses studies which indicate that the Celtic language survived longer than it was previously assumed (see also Trudgill 2011: 51). We can also speak of specific areas of Celtic influence where the importance of northern England is highlighted (see especially, Laker 2002 and 2009 on the phonological impact of Celtic on the Northumbrian dialect of OE, as well as Klemola 2000 and Benskin 2011 on the Northern Subject Rule). In the data of the present study, Celtic influence – if there was some, as suggested by the zone division proposed by White (2002: 154, section 5.4.2) – could be expected to be found predominantly in the North but it would most probably be only reinforcing rather than direct. As will be remembered, the values of the CV, for the SVO order pointed to the most regular distribution in the North, making the dialect stand out in the dialectal spectrum. This could be interpreted as suggesting a combined Celtic and Scandinavian input. Yet, the highest normed frequencies for SVO were unequivocally and repeatedly found in both the North and the East Midlands, the former Danelaw area.

It should also be borne in mind that the Scandinavian impact does not seem to be limited merely to the soon-to-be-caseless nouns and adjectives. Indeed, another factor speaking in favour of the Scandinavian input to the morphosyntactic shift presently discussed is the changing status of pronouns in English. When differentiating nominal and pronominal elements of NPs in the SVO string, the latter appear in the new, no longer clitic, position most often in the two dialects which correspond to the Scandinavian contact areas. The change appears to be both qualitative and quantitative in nature. On one hand, there is a shift from the OE, more pronounced, surface differentiation (Fischer *et al.* 2000: 71), to a growing positional fixity, which aligns with that of nominal constituents (also Howe 1996: 69). On the other, a clear regional preference for SVO with pronominal elements is largely confined to the northernmost dialects. The distribution does not seem to be profoundly affected by the pressures coming from particular clause environments. Nor does it change significantly if the common base is varied or if foreign translations are excluded (Tables 5.1.2.5b-e). The peculiar distribution of pronominal SVO cannot be attributed to the laws of frequency of use either. After all, how many uniform texts types[148] would have to be included in the data samples to produce such telling patterns, based on sequencing which comprises (often more than once) the most frequently used words in Germanic languages?[149] The fact that Old Norse did not treat pronouns as clitics but as full NPs

[148] Taking the set from the North as an example, there are six texts of which three are religious treatises and one is a rule. Together, they fit into the realm of religious/academic prose, which, by default, should be low on pronouns (Biber *et al.* 1998: 69). The remaining two texts are a sermon and a medical handbook. Even if the sermon, dated to the second part of ME, is a classic example of a conversational style - high on pronouns (Biber *et al.* 1999), or as Kohnen (2007) puts it, rambling (2007: 291), the handbook still falls into the category of secular instruction, aligning more with the first four texts. In all, five of six texts that are not pronoun-friendly display particular preference for pronominal SVO.

[149] All the English subjective and objective personal pronouns (except 'us') crop up in the first one hundred most frequent words (Howe 1996: 51). The observation is relevant due to the fact that the choice between particular constituent order in ME seems to have been based on similar aspects to account for linguistic variation in research of contemporary material (e.g. Fischer *et al.* 2000: 174).

has to be, therefore, a key factor here. The changing status of pronouns in English presents itself as a clear example of a convergent development driven by contact with the early Scandinavian population.

Interestingly, the general hierarchy of items affected by language contact puts pronouns at the bottom rather than on the top in the pool of possibilities (e.g. Matras 2009: 133, also Trudgill 2011: 52). However, contact-induced changes involving pronouns do take place. As the early history of English shows, contact influence was already instrumental in pronoun substitution (*they* for the older *hi*, *hie* forms) to resolve a potential lack of clarity between the OE singular and plural paradigms. Thus, it seems plausible enough to ascribe another change in the pronoun system to the same source, especially if the development ultimately achieves the same objective. Admittedly, the impact of French has to be considered as well when analysing the process of decliticisation of pronouns in English. In that case, however, one should expect to find a geographically reverse pattern of distribution, with the southern dialects as its centre. Otherwise, there would be a relatively even spread throughout the dialectal map. As the results show, neither of the two scenarios occurred. By contrast, the distribution of non-pronominal SVO poses some problems at first glance, as a considerable number of instances of non-pronominal SVO ordering are found in the southern rather than the northern dialects. This looks like a countertendency to the layout of pronominal SVO. However, the observed patterns of variation within the SVO strings with non-pronominal objects suggest that the distribution of the feature with the focus placed on nominal components is nothing but a reflection of the syntactic preferences dating to the earliest days of English, when nominal objects tended to be mostly postverbal. The regularity of low CV values for nominal objects in SVO throughout the dialectal spectrum is, indeed, remarkable. The high values of normed frequencies and the absence of sharp fluctuations also seem to support that. It should also be mentioned that the distributional pattern of nominal SVO starts to follow that of the pronominal one, once the common base is increased to 200 clauses. This is a more optimal normalising scale, considering the size of the texts in PPCME2. The highest preference for SVO with nominal objects is again found in the North and the East Midlands at IP-SUB level and when the foreign translations are excluded from calculations. All things considered, the loss of cases has to be considered a relevant factor in the process of syntactic reanalysis. The fact that it does not come out so strongly in the noun-oriented data might mean that the development in question was overshadowed by the scale of changes occurring around the pronominal constituents, especially if these changes originated in the very same dialects.

Finally, to confirm a connection between the convergence point and the effects of contact with the early Scandinavians, one should mention an interesting drop in variation which can be found predominantly in the North and the East Midland sets once the foreign translations are excluded from the samples. The decrease occurs repeatedly within different output configurations in the overall results: within matrix and subordinate clause levels separately, within element strings which include or exclude empty categories, as well as within SVO sequences composed of non-pronominal constituents. The significance of the repeated reduction in variation not only stems from the fact that recurring patterns of distribution of the investigated feature have been

associated with some specific locale corresponding to the relevant contact areas. It is also significant on account of the expected outcomes upon the reduction of the size of data. As observed in the process of handling the output, the usual result of downsizing any data pool created growth in variation (i.e. leading to a higher CV). However, for the northernmost dialects, upon exclusion of translations, the situation was reversed, which clearly shows that pressures on a potential preference (or non-preference) for SVO coming from other languages (with the exception of Scandinavian and Celtic) have to be taken into account. The presence of texts that are based on originals from Latin, French or Dutch weakens geographically determined contrasts in the distribution of SVO.

To conclude, there is more to the Anglo-Scandinavian language contacts than mere lexical additions. As the results of this study have shown, the output alignments meet in the northernmost sectors to such an extent that it is hard to dismiss the significance of changes within NPs altogether, with subsequent consequences on syntax. What is more, the frequent replication of patterns, which cuts the North and the East Midlands off from the rest of the dialect sectors, shows that internal forces, such as drift (section 3.3.1), cannot ultimately explain the regional variation in such an overarching feature as word order.

The third research question tackled in this study has addressed the validity of a theoretical model of convergence, the wave theory, to explain the morphosyntactic changes presently discussed. This theory distinguishes between a focal area, which is the centre of innovation, and the peripheries, which are the adjacent territories, not directly included in the point of origin of changes but which, in turn, absorb (some of) the effects of the innovation. Since the database used in the current research has factored in the two aspects, time and space, which constituted a prism through which changes proceeded in a wave-like fashion, this study aimed to find out if this kind of wave pattern could actually be distinguished in the results. Indeed, the results showed that the repetitive pattern of the highest and most regular (S)VO has pointed specifically to one region, the North. The Scandinavian provenance in making the North a focal area was especially clear when the overall results on the distribution of (S)VO were compared with the results obtained for the subordinate clause level. A higher and more regular distribution of the investigated feature in this dialect agreed with the observations made by Samuels (1985), which portrayed the North as an area more densely populated by both Scandinavian races, hence the deeper linguistic penetration. What also speaks in favour of a stronger impact on English in that part of the country is the fact that the Norse language survived longer north of the Humber. In addition, various sources confirm that the Scandinavian influence was still very much felt during the second part of the Middle English period, despite the royal incentive to stamp out the non-English tendencies lingering in this area (e.g. Bugge 1921: 175).

It is true that the Scandinavians were not alone in having an impact on English culture and eventually language. Apart from Norwegians and Danes residing in the North, there were the Britons and Frisians along with all the refugees and the persecuted from Scotland or the rest of England. Population-wise the North was a multicultural area, where potentially each of these peoples could exert some influence on English. From that standpoint, the highest and most regular preference of (S)VO north

of the Humber revealed itself as proof of substantial external influences on English not exclusively from the Scandinavian side but more collectively, with each population group adding its share. Lastly, the uniqueness of the North must have sprung from its socio-political separatism. On one hand, there was the political autonomy with administrative idiosyncrasies characteristic of the Danelaw; on the other, the unintended but effective isolationism of the northern Church (section 2.3). All these were bound to fuel contrasts. It is therefore justified to regard the North as the hub of linguistic innovation. It was the land where different habits and customs converged and were allowed to thrive.

As we have seen, the height of normed frequencies along with the values of the coefficient of variation, make each of the dialect sets studied here stand apart from one another. Therefore, these sets can be considered to represent a stage of advancement of the change, following McMahon (1994: 228). The analysis of the material from the earlier and later parts of ME, too, showed that the feature was, indeed, more ensconced at the closing of the epoch. Yet, the model's idea of ripples radiating from the centre of innovation to the peripheries does not come out so clearly here, so the applicability of the notion of contagious diffusion is not straightforward. One reason for this is the fact that we are here dealing with the beginning of a long process of syntactic reshufflings, which only ended a few centuries later. Secondly, the emergence of SVO has been a colossal enterprise of various structural rearrangements, which occurred as a result of different kinds of pressures. As for the current study, the distinct distributions of (S)VO at the IP-MAT and IP-SUB levels, too, have clearly shown the effects of the intricate mechanism of syntactic change. Moreover, the further away from the focal area and the more forward in time we move, the greater number of variables come into play and the less confident statements concerning the diffusion of SVO can be made. However, the use of such statistical data handling measures as the CV has allowed for some generalisations to be made.

In general, the wave theory works well in longitudinal comparative studies, with dialect as the unit to fall back on, and where identifying the influential centre is essential. In this study, the model has helped to reveal the repetitive patterns of regional variation in the distribution of (S)VO, as arising particularly from external pressures. On the other hand, the theory seems to be only fairly adequate with respect to depicting the relevant dispersal pattern. Even though the time and space are clearly marked in the data to generate individual values for each of the dialects, one has to take into account the fact that complex developments such as word order entails many secondary processes. Each of these processes may have their own pattern of spread. Yet, provisional peripheries of the wave could be discerned, when contrasting the focal area with the rest of the territories, especially if the influential centre was located in one of the two possible socio-geographic endpoints and if the distribution of the investigated feature within different sample configurations followed similar though not necessarily identical routes.

The last objective of the current study was to assess the usefulness of parsed diachronic corpora such as YCOE and PPCME2 in comparative research on changes within complex syntactic structures. The question to address in this respect was whether the nature and composition of these corpora could be an effective enough

tool for constructing valid arguments concerning the causes of the word order change in early English. Needless to say, the two corpora have been designed explicitly for examining a variety of grammatical features. As the many comparison perspectives described in the results chapter have shown, these databases do, indeed, offer numerous word order oriented parameters to explore. Firstly, YCOE and PPCME2 allowed me to investigate the distribution of different word order arrangements overall as well as at specific clause levels, which was essential to documenting the emergence of prevalent SVO in English. Furthermore, the differentiation between nominal and pronominal elements of NPs, an aspect often neglected in research, was also central to the study of the emergence of the new syntax. In addition, an opportunity to create SVO sequences with, as well as without, empty categories (ECs) revealed that the presence of items not lexically represented in the analysed data, made available in the process of parsing, may considerably change the manner with which scholars make inferences about their investigated feature, word order included. The sequences which contained ECs repeatedly produced distinct and often not so uniform patterns as the sequences which excluded them[150]. Secondly, dialect information assigned to the majority of texts included in the two corpora made it possible to investigate potential external influences, provided the impacts in question were geographically determined (i.e. Celtic and Scandinavian vs. Latin and French). The evidence supporting external pressures was upheld not only on the basis of repeatedly confirmed highest normed frequency values in the sets corresponding to the relevant contact zones, but also by the recurring patterns of pronounced regional variation, as indicated by the values of the CV. Thirdly, the statistic tool of CV enabled gauging the extent of other pressures, associated specifically with the properties of texts included in the corpora. These properties consisted of the MS dating, type of genre and the fact that a given text was a translation from a foreign original, all of which were specified in the databases. Admittedly, the insufficiency of data from some dialect sets somewhat impaired clean, linear comparisons, just as it affected the degree of representativeness of some of the samples. However, the lack of adequate material in the early English corpora is an aspect completely independent of the efforts by both the database compilers and the database users. Furthermore, filtering of results through the coefficient of variation compensated for this shortcoming, notably by allowing repeated sampling. The values of CV in various output set-ups helped to confirm the validity of the findings in spite of the deficiencies in the data. In all, apart from the obvious technical advantage of electronic parsing, which yields quick results from vast amounts of data, YCOE and PPCME2 enable tracking the complex changes from a more all-inclusive perspective. Overarching linguistic features can be investigated thoroughly by simultaneously exploring many different angles, while giving the researcher better control over the material, for instance, by allowing accommodation of specific data retrieval and handling strategies where necessary.

[150] For more considerations around the presence of empty categories see section 7.2.

7.2 CONTRIBUTION TO THE STUDY OF EARLY ENGLISH SYNTAX AND DIRECTIONS FOR FURTHER RESEARCH

The present study has investigated word order change in early English, which was partially triggered and later promoted by external pressures from Scandinavian languages. Although the emergence of SVO as the basic word order has been explored by scholars for decades, no satisfactory explanation has been provided as to the ultimate causes of this change. The controversial aspects of the change revolve mainly around the fact that the English language has diverged radically from the Germanic family matrix, and many of the divergent characteristics were either connected to the changes in word order or they occurred alongside them (e.g. McWhorter 2002). Section 4.4.2 outlined some of the most influential contributions to the topic. This research offers yet another perspective on the origins of the emerging new syntax in early English. The research has been designed to track the change from the end-result backwards in time. In other words, this study has focused on an analysis of the full surface SVO and VO sequences, i.e. the patterns used in present-day English, in texts from the early formative periods. Apart from the geographically determined Scandinavian impact, other influencing factors such as archaeology and human genetics have also been explored (sections 2.1-2). This approach, I would argue, has opened up important new perspectives on the radical changes that the English language underwent in these periods.

Firstly, the current research has highlighted the importance of language contacts as a motive for linguistic change, thus answering on its part to the challenge presented by many linguists, who claim that the role of external influences on the history of English has been underestimated in previous studies (e.g. Hickey 2010: 4). The prevailing view on the nature and outcomes of the Anglo-Scandinavian language contacts has been that of the external pressures acting as a mere intensifier of changes, rather than as a driving force. What is more, the scope of such influences has been largely limited to lexis, as pointed out for example by Trips (2002: 332), and any claims concerning contact influences in morphosyntax have been either refuted or only 'tentatively' agreed on (e.g. Thomason and Kaufman 1988: 302-315). The present study has shown that the clear regional preference for the distribution of (S)VO with the new pronominal layout converged with the morphological simplification exactly in the dialect sets which corresponded to the areas where the Anglo-Scandinavian impact was most diverse and most pronounced. These areas were the same territories which were visited by the early North Germanic people long before the advent of the Viking Era[151] as well as during the raid of the northern marauders. These same lands were later densely peopled by the early Scandinavians. In due course, as both the historical and linguistic evidence shows, following the general assimilation of the North Germanic population into English society, the Nordic footprint was long-lasting there, apparently surviving the ethnic cleansing carried out on the orders of William I. The role of the North, in particular, should be mentioned here, as the 'fifth part of

[151] As confirmed by numerous archaeological and genetic studies (section 2.1-2)

England' held not only the Scandinavian populations, i.e. Norwegians and Danes, but included the already indigenous Celts. The multitude of peoples must have reinforced the pressures in the early English language which eliminated contrasts (i.e. loss of inflections) and promoted similarities (i.e. pronouns as full NPs and more opportunities to lean on (S)VO). The socio-political conditions of the North, especially with its separatism, played a role as well. It becomes clearer, therefore, why the results from the southern half of the Scandinavian zone did not align so neatly with those from the North. The southern half was mostly peopled by the Anglo-Danish population, lacking the additional Norwegian and Celtic input and expressing socio-politically less subversive sentiments. Lastly, the regional preference for the new syntax in a specific locale not only moves internal pressures, such as drift, more into the background as an explanatory factor but also makes one wonder how much syntactically different English would be from the other Germanic languages had it not been for these contacts (e.g. McWhorter 2002: 217-272, with a revised view 2009: 163-191, 2011: 262), and whether SVO would have become a primary marker of grammatical functions at all. Interestingly, the arguments presented here overlap, to some extent, with the ideas put forward by Trips (2002), presented in chapter 4. The importance of the recognition of the Danelaw area is emphasised in both studies. However, the extent of the Scandinavian impact presently advocated takes into account the internal pressures affecting the English language prior to this contact. The effects of this interaction with the early Scandinavian population are perceived as the force driving on the changes which began before the contact in question.

Secondly, the present study has stressed the role of convergence models when examining the contact induced morphosyntactic changes in early English through the use of parsed diachronic corpora. The selected model has been the wave theory, which, since its introduction to the pool of theoretical possibilities, has long since been followed by other, often more factor-sensitive frameworks. Nevertheless, the wave theory has proved to be an adequate model not only due to the fact that its potential revolves around the two important aspects marked in the databases, both of which are relevant to the scope of this study, namely time and space. This research has also shown that the wave model can be used for tracking complex, overarching changes, such as word order shifts, provided that some additional data managing tools are used to both check and strengthen the arguments. The multiple sampling strategies backed by the statistical tool of coefficient of variation has enabled identifying the focal area, the centre of innovation from which the novel syntactic preferences have radiated to other territories of the English-speaking realm. Admittedly, the theory falls slightly short with respect to showing a neat dispersal pattern of the investigated feature. Yet, this shortcoming is determined more by the nature of the change itself rather than by the simplistic character of the model. Whichever model is chosen, it would be difficult to fully explain the complexity of the spread of the new syntactic configuration, as this configuration itself subsequently involved a number of distinct accommodating developments once the innovation had left the centre. Finally, the current study has sought to address the call made by Renfrew (1989), in which he notes that the explanatory power of the convergence models has not been exhaustively spent in investigations of language change viewed over longer periods of time. In addition, Renfrew

suggests that these models be used more often in historical linguistics to account for the intricacies of language change (Renfrew 1989: 112).

Thirdly, this research has hopefully introduced a new angle when it comes to methodology within the field of early English historical syntax, trying to answer the demand for an extensive quantitative approach (e.g. Fischer *et al.* 2000: 160) and for "more expedient sampling methods" (Biber *et al.* 1998: 253) in word order studies. Relying on the input provided by the early English corpora such as YCOE and PPCME2, the current research has focused on sequences of SVO and VO constituents, retrieved from a wide array of data samples. Each of these samples has been created to account for a particular aspect relevant to determining the range of the distribution of the investigated feature. One of those aspects has been a given clause environment, where the new (S)VO would manifest differently, being partially regulated by the earlier syntactic conditions prevailing in each clause. Similarly, comparisons have been made between the behaviour of nominal and pronominal constituents in word order sequences.

Data-specific factors explored in this study included some characteristics of early English texts, such as the distinction between earlier renditions and later copies, as well as the potential impact from foreign translations. An attempt to evaluate the influence of different genres, though less fruitful, was also made. As the results showed, each of these factors had a bearing on the distribution of (S)VO in a particular manner. Ultimately, the dialect factor turned out to be a decisive determinant of the prevalence of the feature in early English, as confirmed by the comparison of text doublets (i.e. dialect translations). Furthermore, several data handling techniques were employed to subsequently elicit the relevant outcomes, among them the use of the coefficient of variation and the multiple common base approach. The former accounted for how distinct data sets arrived at average (normed) frequency values; the latter allowed examining the stability of the results measured against corpora comprised of different text sizes.

Finally, just as the present study has focused on the analysis of various conditions affecting the distribution of early English word order, it has also relied on input from other disciplines to support and corroborate the findings from linguistic data. The fragment of Higden's *Polychronicon*, through the hand of Trevisa and quoted in section 2.4, was one of the notable fourteenth-century instances, already showing how the early English writers were conscious of the extent to which their language had been affected by the country's history. It seemed only logical, therefore, to add as much relevant information as possible on the land and its inhabitants, including the evidence from archaeology (e.g. Miles 2006) and population genetic studies (e.g. Oppenheimer 2006, Sykes 2007). This interdisciplinary approach has helped to achieve a more accurate interpretation of the numerical data obtained from the diachronic corpora.

In conclusion, to say that the study of the emergence of SVO in English offers several possible angles for research would amount to an understatement. The emergence of strict SVO in the English language provides an almost infinite number of possibilities, depending on the scope of investigation and the methods used, especially in view of the growing interest in the use of more sensitive statistical tools for handling language data. Historical linguists have been well aware of the investigative potential

behind this morphosyntactic change and, indeed, numerous interesting solutions and explanations have been provided over the decades. As regards the current study, every factor determining the layout of the investigated feature that has been mentioned could be further expanded or refined by employing alternative sampling strategies or other theoretical models.

One aspect that could be explored further is the presence of empty categories in the parsing scheme and their role in forming arguments on word order distributions in English. In the present study, empty categories were handled collectively in order to test the applicability of the parsed diachronic corpora, by indicating differences in the frequency of the distribution of the investigated feature when they were included or excluded from searches. Subsequent analysis could focus on particular types of these categories. No doubt some of them would offer the opportunity of comparing syntactic environments between the earlier and later stages of development of English. Empty expletive subjects and other, "small-pro", i.e. pro-drop subjects are the relevant instances (cf. Fischer *et al.* 2000: 39, 71). Others present an interesting direction for research due to the syntactic setting to which they are limited, e.g. arbitrary PRO subjects which occur in the exceptional case-marking (ECM) constructions[152] (for details see e.g. Fischer *et al.* 2000: 222-5). Furthermore, greater attention could also be devoted to the influence of genres, which, like other technical hurdles encountered when dealing with diachronic corpora of Old and Middle English, presents a problem because of the shortage of adequate text specimens. My results have shown that despite the division of data into larger prototypical text categories, the influence of other factors on word order distribution turned out to be more important than the impact of genres. It remains to be seen whether further investigation of the influence of genres would change the picture emerging from this study.

[152] IP-INF were not investigated in the present study.

Bibliography

Abbott, M. 1991. 'Macushi'. In Derbyshire, D. C. and Pullum, G. K. (eds) *Handbook of Amazonian Languages 3*. Berlin: Mouton de Gruyter, 23-160.

Ahlqvist, A. 2010. 'Early English and Celtic'. *The Australian Celtic Journal,* 9: 43-73.

Aitchison, J. 2003. 'Metaphors, Models and Language Change'. In Hickey, R. (ed.), 39-53.

Allen, C. 2002. 'Case and Middle English Genitive NPs'. In Lightfoot, D. W. (ed.), 57-87.

Allen, C. L. 1995. *Case Marking and Reanalysis: Grammatical Relations from Old to Early Modern English*. Oxford: Clarendon Press.

Allen, C. L. 2000. 'Obsolescence and sudden death in syntax: The decline of verb-final order in early English'. In Bermudez-Otero, R., Denison D. and Hogg, R. (eds), *Topics in English Linguistics (TIEL): Generative Theory and Corpus Studies: A Dialogue from 10 ICEHL*, Munchen: Walter de Gruyter, 3-25.

Allen, C. L. 2006. 'Case Syncretism and Word Order Change'. In van Kemenade, A. and Los, B. (eds), 201-223.

Andersen, P. K. 1983. *Current Issues in Linguistic Theory, Volume 25: Word Order Typology and Comparative Constructions*. Philadelphia, PA: John Benjamins Publishing.

Anderson, J. 1986. 'A note on Old English impersonals'. *Linguistics* 22: 167– 177.

Anderson, J. 1993. 'Parameters of syntactic change: a notional view'. In Jones, Ch. (ed.), *Historical Linguistics: Problems and Perspectives*. New York: Longman, 1-42.

Auwera, J. van der and Genee, I. 2002. 'English do: On the convergence of languages and linguists'. *English Language and Linguistics* 6.2: 283-307.

Bailey, C. J. N. and Maroldt. K. 1977. 'The French Lineage of English'. In Meisel, J. (ed.), *Langues en contact–Pidgins–Creoles–Languages in Contact.* Tübingen: Gunter Narr, 21-53.

Bailey, Ch. J. N. 1973. *Variation and Linguistic Theory*. Arlington: the Center for Applied Linguistics.

Ball, C. 1991. *The historical development of it-cleft.* University of Pennsylvania Dissertation in Linguistics, Ann Arbor, MI: University Microfilms International.

Balsley, H. L. 1964. *Introduction to statistical method.* Totowa, N.J.: Littlefield, Adams and Co.

Barlow, F. 1966. 'The Effects of the Norman Conquest'. In Whitelock et al., 123-164.

Barnes, M. 1993. 'Norse in the British Isles'. In Faulkes, A. and Perkins, R. (eds), *Viking Revaluations: Viking Society Centenary Symposium*. London: University College London, 65-84.

Battye, A. and Roberts, I. (eds). 1995. *Clause structure and language change*. Oxford: Oxford University Press.

Bean, M. C. 1983. *The Development of Word Order Patterns in Old English*. London: Croom Helm.

Bech, K. 2001. *Word order patterns in Old and Middle English: a syntactic and pragmatic study*. Ph.D dissertation.

Bense, J. F. 1925. *Anglo-Dutch Relations from the Earliest Times to the Death of William the Third*. The Hague: Martinus Nijhoff.

Benskin, M. 2011. 'Present indicative plural concord in Brittonic and early English'. *Transactions of the Philological Society* 109: 158–85.

Bentz, Ch. and Christiansen, M. H. 2013. 'Linguistic adaptation: The trade-off between case marking and fixed word orders in Germanic and Romance languages'. In Peng, G. and Shi, F. (eds), *Eastward Flows the Great River: Festschrift in Honor of Prof. William S-Y. Wang on his 80th Birthday*. Hong Kong: City University of Hong Kong Press, 45-61.

Bertacca, A. 2009. *Natural Morphology and the Loss of Nominal Inflections in English*. Pisa: Edizioni Plus – Pisa University Press.

Biber, D., Conrad, S. and Reppen, R. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.

Bibre, P. 2001. 'North Sea Language Contacts in the Early Middle Ages: English and Norse'. In Liszka, T. R. and Walker, L. E. M. (eds), *The North Sea World in the Middle Ages: Studies in the Cultural History of North-Western Europe*. Dublin: Four Courts, 88-107.

Blair, P. H. 1977. *An Introduction to Anglo-Saxon England*. Cambridge: Cambridge University Press.

Blair, P. H. 1984. *Anglo-Saxon Northumbria*. London: Variorum Reprints.

Blake, B. J. 2004. *Case*. Cambridge: Cambridge University Press.

Blake, N. (ed.). 1992. *The Cambridge History of the English Language: Volume II 1066-1476*. Cambridge: Cambridge University Press.

Blythe, R.A. and Croft, W. 2012. 'S-curves and the mechanisms of propagation in language change'. *Language* 88, 2: 269-304.

Bobaljik, J. and Thráinsson, H. 1998. 'Two heads aren't always better than one'. *Syntax* 1. 37-71.

Braaten, B. 1967. 'Notes on continuous tenses in English'. *Norsk Tidsskrift for Sprogvidenskap* XXI: 167–80.

Britton, D. 1991. 'On Middle English *she, sho*: A Scots Solution to an English Problem'. *Nowele* 17: 3-51.

Brook, G.L. 1972. *English Dialects*. London: Deutsch.

Bugge, A. 1921. 'The Norse settlements in the British Islands'. *Transactions of the Royal Historical Society*, 4th Series, 173-210.

Burnley, D. 1992. 'Lexis and semantics'. In Blake (ed.), 409-499.

Butler, Ch. S. 1985. *Statistics in Linguistics*. Blackwell Publishers.

Bybee, J. I., Perkins, R. and Pagulica, W. 1994. *The evolution of grammar: tense, aspect and modality in the languages of the world*. Chicago Il.: The University of Chicago Press.

Bynon, T. 1977. *Historical Linguistics*. Cambridge: Cambridge University Press.

Calmorin, L. 1997. *Statistics in Education and the Sciences*. Manila: Rex Book Store Inc.

Campbell, L. 1981. 'Generative phonology vs. Finnish Phonology: Retrospect and Prospect'. In Goyvearts, D. L. (ed.), *Phonology in the 1980's*. Ghent: Story-Scientia, 147-182.

Campbell, L. 1997. 'Amerindian personal pronouns: a second opinion'. *Language* 73: 339-51.

Carlier, A. and Verstraete, J-Ch. 2013. *The Genitive* (Case and Grammatical Relations across Languages, vol. 5). Amsterdam: John Benjamins Publishing.

Carus-Wilson, E. M. 1958. 'Towns and trade'. In Poole, A. L. (ed.), 209-263.

Clackson, J. and Horrocks, G. 2007. *The Blackwell History of the Latin Language*. Malden: Blackwell Publishing.

Claudi, U. 1994. 'Word order change as category change'. In Pagliuca, W. (ed.), 191– 231.

Cloutier, R. A. 2005. 'Review of Carola Trips' From OV to VO in Early Middle English'. *English Language and Linguistics* 9.1: 181-191.

Collingwood, W.G. 1993. *Scandinavian Britain*. Felinfach: Llanerch Publishers.

Colman, F. 1988. 'Heavy arguments in Old English'. In Anderson, J. M. and Macleod, N. (eds), *Edinburg Studies in the English Language*, 1. John Donald: Edinburgh (1988), 33-89.

Colman, F. (ed.). 1992. *Evidence for Old English: Material and Theoretical Bases for Reconstruction.* (Edinburgh Studies in the English Language, 2) John Donald, Edinburgh.

Corrie, M. 2006. 'Middle English – Dialects and Diversity'. In Mugglestone, L. (ed.), 86-119.

Cowell, M. W. 1964. *A Reference Grammar of Syrian Arabic (based on the dialect of Damascus)*. Washington: Georgetown University Press.

Cramp, R. 1982. 'The Viking Image'. In Farrell, R. T. (ed.), 8-19.

Crawford, B. E. 2003. 'The Vikings'. In Davies, W. (ed.) *From the Vikings to the Normans*. Oxford: Oxford University Press, 41-72.

Cristofaro, S. 2011. 'Language Universals and Linguistic Knowledge'. In J. J. Song (ed.), 227-249.

Croft, W. 2000. *Explaining Language Change: An Evolutionary Approach*. Harlow: Pearson Education Limited.

Crowley, M. J. 2007. *The R Book*. Hoboken: Wiley.

Crystal, D. 2003. *The Cambridge Encyclopaedia of the English Language*. Cambridge: Cambridge University Press.

Cunliffe, B. 2004. *Facing the Ocean: the Atlantic and Its Peoples*. Oxford: Oxford University Press.

Curzan, A. 2003. *Gender Shifts in the History of English*. Cambridge: Cambridge University Press.

Czerniak, I. 2011. 'Anglo-Scandinavian language contacts and word order shift in early English'. In Fisiak, J. and Bator, M. (eds), *Foreign Influences on Medieval English* (Studies in English Medieval Language and Literature vol. 28). Frankfurt am Main: Peter Lang, 139–154.

Dal, I. 1952. 'Zur Entstehung des englischen Participium Praesentis auf –ing'. *Norsk Tidsskrift for Sprogvidenskap* XVI: 5–116.

Dance, R. 2012. 'English in Contact: Norse'. In Bergs, A. and Brinton, L. J. (eds), *English Historical linguistics: An International Handbook, Vol II* (Handbooks of Linguistics and Communication Science 34.2) Berlin and New York: De Gruyter, 1724-1737.

Danchev, A. 1991. 'Language change typology and some aspects of the SVO development in English'. In Kastovsky, D. (ed.), *Historical English Syntax*. Berlin: Mouton de Gruyter, 103-124.

Davies, W. (ed.). 2003. *From the Vikings to the Normans*. Oxford: Oxford University Press.

Davis, G. 2006. *Comparative Syntax of Old English and Old Icelandic*. Studies in Historical Linguistics, Vol.1. Bern: Peter Lang.

Dawson, H. C. 2003. 'Defining the outcome of language contact: Old English and Old Norse'. *Ohio State University Working Papers in Linguistics* 57: 40–57.

De Vogelaer, G. 2007. 'Extending Hawkins' comparative typology: case, word order, and verb agreement in the Germanic languages'. *Nordlyd: Tromso University Working Papers on Language and Linguistics*, 34: 167-182. Available at http://septentrio.uit.no/index.php/nordlyd/article/view/93, accessed April 11, 2011.

Dekeyser, X. 1986. 'English contact clauses revisited: a diachronic approach'. *Folia Linguistica Historica* 7.1: 107–20.

Denison, D. 1993. *English Historical Syntax*. New York: Longman Publishing.

Deutscher, G. 2000. 'Stability against the odds? The survival of verb final in Akkadian'. In Sornicola, R., Poppe, E. and Shisha Halevy, A. (eds), *Stability, Variation and Change of Word-Order Patterns over Time*. Amsterdam: John Benjamins, 55-69.

Diensberg, B. 1996. 'French transplanted: A re-evaluation of the importance of Anglo-French in the developments of the English language during the Middle English period'. In Fisiak, J. (ed.), *Middle English miscellany: From vocabulary to linguistic variation*. Poznań: MOTIVEX, 253-266.

Douglas, D. C. 1966. 'William the Conqueror: Duke and King'. In Whitelock, D. et al., 45-76.

Downes, W. 1998. *Language and Society*. Cambridge: Cambridge University Press.

Dryer, M. S. 1991. 'SVO languages and the OV:VO typology'. *JL* 27: 443-82.

Dryer, M. S. 1997. 'On the six-way word order typology'. *SiL* 21: 69-103.

Dryer, M. S. 2013a. 'Order of Subject, Object and Verb'. In Dryer, M. S. and Haspelmath, M. (eds), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available at http://wals.info/chapter/81, accessed Nov 21, 2014.

Dryer, M. S. 2013b. 'Determining Dominant Word Order'. In Dryer, M. S. and Haspelmath, M. (eds), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available at http://wals.info/chapter/s6, accessed Nov 21, 2014.

Emonds, J. E. and Faarlund, J. T. 2014. *English: The language of the Vikings*. Olomouc: Palacký University.

Enrique-Arias, A. 2010. 'On language contact as an inhibitor of language change'. In Breitbarth, A., Lucas, Ch., Watts, Sh., and Willis, D. (eds), *Continuity and Change in Grammar*. Amsterdam: John Benjamins Publishing Company, 97-118.

Faarlund, J. T. 2001. 'From Ancient Germanic to modern Germanic languages: Diachronic aspects of language types and linguistic universals'. In Haspelmath, M., König, E., Oesterreicher, W., and Raible, W. (eds), *Language Typology and Language Universals: An International Handbook*, Vol. 2 [HSK 20(1)], Berlin: Mouton de Gruyter, 1706-1719.

Faarlund, J. T. 2004. *The Syntax of Old Norse*. Oxford: Oxford University Press.

Faarlund, J. T. 2008. 'A mentalist interpretation of grammaticalization theory'. In Eythorson, T. (ed.), *Grammatical Change and Linguistic Theory: The Rosendal Papers*. Amsterdam: John Benjamins Publishing Company, 221-244.

Farrar, K. and Jones, M.C. 2002. 'Introduction'. In Jones, M. C. and Esch, E. (eds), *Contributions to the Sociology of Language [CSL]: Language Change: The Interplay of Internal, External and Extra-linguistic Factors*. Berlin: Mouton de Gruyter, 1-18.

Farrell, R. T. (ed.). 1982a. *The Vikings*. London: Fillimore & Co. Ltd.

Fellows-Jensen, G. 1991. 'Scandinavian Influence on the Place-names of England'. In Ureland, P.S. and Broderick, G. (eds), *Language Contact in the British Isles. Proceedings of the Eighth International Symposium on Language Contact in Europe, Douglas, Isle of Man, 1988*. Tubingen: Max Niemeyer Verlag, 337-354.

Fernandez-Cuesta, J. and M. N. Rodriguez-Ledesma. 2006. 'Northern Middle English'. In Dury, R., Dosena, M. and Gotti, M. (eds), *English Historical Linguistics 2006: Selected papers From the Fourteenth International Conference on English Historical Linguistics (ICEHL 14), Bergamo, 21-25 August 2006*. Amsterdam: John Benjamins Publishing, 91-109.

Filppula, M. 2003. 'More on the English progressive and the Celtic connection'. In Tristram, H. L. C. (ed.), 150-168.

Filppula, M. 2009. 'The rise of it-clefting in English: areal-typological and contact-linguistic considerations'. In Filppula, M. and Klemola, J. (eds), *Special issue on Re-evaluating the Celtic hypothesis – English Language and Linguistics* 13.2: 267-293.

Filppula, M. 2010. 'Contact and the Early History of English'. In Hickey, R. (ed.), 432-453.

Filppula, M., Klemola, J. and H. Pitkänen (eds). 2002. *The Celtic Roots of English* (Studies in Languages 37). Joensuu: University of Joensuu, Joensuun Yliopistopaino.

Filppula, M., Klemola, J. and Paulasto, H. 2008. *English and Celtic in contact*. London and New York: Routledge.

Fischer O. and van der Wurff, W. 2006. 'Syntax'. In Hogg, R. and Denison, D. (eds), 109-198.

Fischer, O. 1992. 'Syntax'. In Blake, N. (ed.), *The Cambridge History of The English Language: Volume II 1066-1476*. Cambridge: Cambridge University Press, 207-398.

Fischer, O. 1994. 'The fortunes of the Latin-type accusative and infinitive construction in Dutch and English compared'. In Swan et al. (eds), 91-134.

Fischer, O., van Kemenade, A., Koopman, W. and van der Wurff, W. 2000. *The Syntax of Early English*. Cambridge: Cambridge University Press.

Fischer, S. 2010. *Word-Order Change as a Source of Grammaticalisation*. Amsterdam: John Benjamins Publishing Company.

Fisher, D. J. V. 1973. *The Anglo-Saxon Age c. 400–1042*. New York: Barnes and Noble Books.

Fisiak, J. 1977. 'Sociolinguistic and Middle English: Some Socially Motivated Changes in the History of English'. *Kwartalnik Neofilologiczny*, 24: 247-59.

Fisiak, J. 2004. *A Short Grammar of Middle English*. Poznań: Wydawnictwo Poznańskie.

Fisiak, J. 2005. *An Outline of English: Volume One – External History*. Poznań: Wydawnictwo Poznańskie.

Fisiak, J. (ed.). 1981. *Historical Syntax* (Trends in Linguistics: Studies and Monographs 23). Berlin: Mouton Publishers.

Foster, S. M. 2004. *Picts, Gaels and Scots*. Batsford: Historic Scotland.

François, A. 2014. 'Trees, Waves and Linkages: Models of Language Diversification'. In Bowern, C. and Evans, B. (eds), *The Routledge Handbook of Historical Linguistics*. Routledge, 161-189, 2014, 978-0-41552-789-7. <halshs-00874726v2>. Available at https://halshs.archives-ouvertes.fr/halshs-00874726v2, accessed April 16, 2015.

Frank H. and Althoen, S. C. 1994. *Statistics: concepts and applications*. Cambridge: Cambridge University Press.

Frank, R. 1989. 'Did Anglo-Saxon audiences have a skaldic tooth?'. In Niles, J.D. and Amodio, M. (eds), *Anglo-Scandinavian England: Norse-English Relations in the Period before the Conquest*. Lanham: University Press of America, Inc., 53-68.

Fuss, E. 1998. *Zur Diachronie von Verbzweit. Die Entwicklung von Verbstellungsvarianten im Deutschen und Englischen*. M. A. Thesis Johann Wolfgang Goethe-Universität: Frankfurt.

Geipel, J. 1971. *The Viking Legacy: The Scandinavian Influence on the English and Gaelic Languages*. London: David and Charles Publishers Limited.

Gelderen, E. van. 2004. *Grammaticalization as Economy* [Linguistik Aktuell/Linguistics Today 71]. Amsterdam: John Benjamins.

Gerritsen, M. 1984. 'Divergent word order developments in Germanic: A description and a tentative explanation'. In Fisiak, J. (ed.), *Historical Syntax*. Berlin: Mouton, 107-136.

Gillingham, J. 2003. 'Britain, Ireland, and the South'. In W. Davies. (ed.), 203-232.

Givón, T. 1979a. 'From discourse to syntax: Grammar as a processing strategy'. In Givón, T. (ed.), *Syntax and Semantics, Vol. 1: Discourse and Syntax*. New York NY: Academic Press, 81– 112.

Givón, T. 1979b. *On Understanding Grammar*. NY: Academic Press.

Godden, M., Gray, D. and Hoad, T. (eds). 1994. *From Anglo-Saxon to Early Middle English: Studies Presented to E. G. Stanley*. Oxford: Clarendon.

Godfrey, J. 1962. *The Church in Anglo-Saxon England*. Cambridge: Cambridge University Press.

Gordon, E. V. 1981. *An Introduction to Old Norse, Second Edition*. Oxford: Oxford University Press. [First Edition published in 1927]

Görlach, M. 1997. *The Linguistic History of English*. London: McMillan Press Ltd.

Greenberg, J. H. 1963b. 'Some universals of grammar with particular reference to the order of meaningful elements'. In Greenberg, J. H. (ed.), 58-90.

Greenberg, J. H. (ed.). 1963a. *Universals of Language*. MIT Press: Cambridge, MA.

Haeberli, E. 1999. *Features, categories and the syntax of A-positions. Synchronic and diachronic variation in the Germanic languages*. Doctoral dissertation, University of Geneva.

Haeberli, E. 2001. 'Speculations on the syntax of subordinate clauses in OE'. *Reading Working Papers in Linguistics* 5: 201-229.

Haeberli, E. 2002. 'Inflectional morphology and the loss of V2 in English'. In Lightfoot, D. (ed.), 88-106.

Haeberli, E. 2010. 'Investigating Anglo-Norman Influence on Late Middle English Syntax'. In Ingham, R. (ed.), *The Anglo-Norman Language and Its Context*. Woodbridge: Boydell and Brewer, 143-63.

Hall, C. 1960. *A Concise Anglo-Saxon Dictionary*, Fourth Edition. Cambridge: Cambridge University Press.

Hamp, E. P. 1975-76. 'Miscellania Celtica I, II, III, IV'. *Studia Celtica* 10/11: 54-73.

Harris, M. B. 1981. 'On the strengths and weaknesses of a typological approach to historical syntax'. In Fisiak, J. (ed.), 183-198.

Haspelmath, M. 2004. 'On directionality in language change with particular reference to grammaticalization'. In Fischer, O., Norde, M. and Perridon, H. (eds), *Up and Down the Cline*. Amsterdam: John Benjamins, 17-44.

Hawkins, J. A. 1986. *A Comparative Typology of English and German: Unifying the Contrasts*. London: Croom Helm.

Heath, J. 1984. *A Functional Grammar of Nunggubuyu*. Atlantic Highlands N. J. / Canberra: Humanities Press / Australian Institute of Aboriginal Studies.

Heine, B. 1994. 'Grammaticalization as an Explanatory Parameter'. In Pagliuca, W. (ed.), 254-287.

Hendricks, J. 2012. 'English in Contact: German and Dutch'. In Bergs, A. and Brinton, L. J. (eds), *English Historical linguistics: An International Handbook, Vol II* (Handbooks of Linguistics and Communication Science 34.2) Berlin and New York: De Gruyter, 1659-1670.

Hiatt, A. 2007. 'Genre without System'. In P. Strohm (ed.), 277-294.

Hickey, R. 1995. 'Early contact and parallels between English and Celtic'. VIEW[Z] (*Vienna English Working Papers*) 4.2: 87-119.

Hickey, R. 2002. 'Internal and external forces again: changes in word order in Old English and Old Irish'. *Language Sciences* 24: 261–283.

Hickey, R. 2012. 'Internally- and Externally- Motivated Language Change'. In Hernández-Campoy, J. M. and Conde-Silvestre, J. C. (eds), *The Handbook of Historical Sociolinguistics*. Malden: Wiley-Blackwell, 387-407.

Hickey, R. (ed.). 2003. *Motives for Language Change*. Cambridge: Cambridge University Press.

Hickey, R. 2012. 'Early English and the Celtic Hypothesis'. In Nevalainen, T. and Traugott, E. C. (eds), *Oxford Handbook of the History of English*. Oxford: Oxford University Press, 497-507.

Hickey, R. (ed.). 2010. *The Handbook of Language Contact*. Malden, MA: Wiley-Blackwell.

Higham, N. J. 1993. *The Kingdom of Northumbria AD 350-1100*. Stroud: Alan Sutton Publishing Limited.

Higham, N. J. (ed.). 2007. *Britons in Anglo-Saxon England*. Woodbridge: Boydell & Brewer.

Hiltunen, R. 1983. *The Decline of the Prefixes and the Beginnings of the English Phrasal Verb* (Annales Universitatis Turkuensis, Ser. B., Tom. 160). Turku: Turun Yliopisto.

Hines, J. 1984. *The Scandinavian Character of Anglian England in the pre-Viking Period*. Oxford: B.A.R. (British Series 124).

Hinterhölzl, R. and Petrova, S. (eds). 2009. *Information Structure and Language Change: New Approaches to Word Order Variation in Germanic* (Trends in Linguistics: Studies and Monographs 203). Berlin: Mouton de Gruyter.

Hoad, T. 2006. 'Preliminaries: Before English'. In Mugglestone, L. (ed.), 7-31.

Hock, H. H. and Joseph, B. D. 1996. *Language History, Language Change and Language Relationship. An Introduction to Historical and Comparative Linguistics*. (Trends in Linguistics. Studies and Monographs 93). Berlin and New York: Mouton de Gruyter.

Hock, H. H. 1986. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.

Hogg, R and Denison, D. (eds). 2006a. *A History of the English Language*. Cambridge: Cambridge University Press.

Hogg, R. and Denison, D. 2006b. 'Overview'. In Hogg, R. and Denison, D (eds), 1-42.

Hogg, R. 1997. 'Some remarks on case marking in Old English'. *Transactions of the Philological Society*, Volume 95:1, 95-109.

Hogg, R. (ed.). 1992. *The Cambridge History of the English Language I*. Cambridge: Cambridge University Press.

Holm, J. 2010. 'Contact and Change: Pidgins and Creoles'. In Hickey, R. (ed.), 252-262.

Holmberg, A. and Platzack, Ch. 1995. *The role of Inflection in Scandinavian syntax*. Oxford: Oxford University Press.

Hopper, P. J. and Traugott, E. C. 2003. *Grammaticalization* (2nd ed.). Cambridge: CUP.

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

Horobin, S. and Smith, J. 2002. *An Introduction to Middle English*. Edinburgh University Press.

Howe, S. 1996. *Study of Personal Pronouns in the Germanic Languages: A Study of Personal Pronoun Morphology and Change in the Germanic Languages from the First Records to the Present Day*. Berlin & New York: Walter de Gruyter.

Hróarsdóttir, T. 2000. *Word Order Change in Icelandic: from OV to VO*. Amsterdam: John Benjamins.

Hróarsdóttir, T. 2009. 'OV langauges: Expression of Cues'. In Hinterhölzl, R. and Petrova, S. (eds), 67-90.

Hudson, R. A. 1980. *Sociolinguistics*. New York: Cambridge University Press.

Hughes, G. 2000. *A History of English Words*. Oxford: Blackwell.

Iglesias–Rábade, L. 2003. *Handbook of Middle English: Grammar and Texts*. (LINCOM Studies in English Linguistics 05). Muenchen: LINCOM GmbH.

Ingham, R. 2009. 'Middle English and Anlgo-Norman in Contact'. In *English Literary Society of Japan Conference*, Tokyo University, May 2009.

Irvine, S. 2006. 'Beginnings and Transitions: Old English'. In L. Mugglestone (ed.), 32-60.

Jespersen, O. 1935. *Growth and Structure of the English Language*. Leipzig: B. G. Teubner.

Joseph, B. D. and Janda, R. D. (eds). 2003. *The Handbook of Historical linguistics*. Malden: Blackwell Publishing.

Jucker, A. H. 1990. 'Word order changes in early Middle English: Some evidence against the conservatism of subordinate clauses'. *Studia Anglica Poznaniensia* 23: 31-42.

Kastovsky, D. 1992. 'Typological reorientation as a result of level interaction: the case of English morphology'. In Kellermann, G. and Morrissey, M.D. (eds), *Diachrony within synchrony: Language history and cognition*. (Dusiburger Arbeiten zur Sprach- und Kulturwissenschaft 14). Frankfurt/M.: Lang, 411-428.

Kastovsky, D. 1994. 'Typological differences between English and German morphology and their causes'. In Swan et al. (eds), 135-157.

Kastovsky, D. 2006. 'Vocabulary'. In Hogg, R. and Denison, D. (eds), 199–270.

Kastovsky, D. (ed.). 1991. *Historical English Syntax*. Berlin: Mouton de Gruyter.

Keller, R. 1994. *On Language Change: The Invisible Hand in Language*. Routledge. [translated by B. Nerlich].

Keller, W. 1925. 'Keltisches im englischen Verbum'. *Anglica: Untersuchungen zur englischen Philologie = Festschrift für Aloys Brandl,* vol. 1. Leipzig: Mayer and Müller, 55-66.

Kemenade, A. van 1987. *Syntactic Case and Morphological Case in the History of English*. (PhD dissertation, Utrecht.) Foris, Dordrecht.

Kemenade, A. van and Los, B. 2006. *The Handbook of the History of English*. Malden: Blackwell Publishing.

Kemenade, A. van and Vincent, N. (eds). 1997. *Parameters of morphosyntactic change*. Cambridge: Cambridge University Press.

Ker, N. R. 1965. *Catalogue of Manuscripts Containing Anglo-Saxon*. Oxford: Oxford University Press.

Kiparsky, P. 1988. 'Phonological change'. In Newmeyer, F.J. (ed.), *Linguistics: the Cambridge survey, vo.1: Linguistic theory: foundations*. Cambridge: Cambridge University Press, 363-415.

Kiparsky, P. 1997. 'The rise of positional licensing'. In van Kemenade, A. and Vicent, N. (eds), 460-494.

Kiparsky, P. 2008. 'Universals Constrain Change; Change Results in Typological Generalizations'. In Good, J. (ed.), *Language Universals and Language Change*. Oxford: Oxford University Press, 23-53.

Klemola, J. 2000. 'The origins of the Northern Subject Rule: A case of early contact?'. In Tristram H.L.C. (ed.), 329-346.

Klemola, J. 2002. 'Periphrastic DO: Dialectal distribution and origins'. In Filppula et al. (eds), 199-210.

Knowles, G. 1997. *A Cultural History of the English Language*. London: Arnold.

Kohnen, T. 2007. 'Connective profiles' in the history of English texts: Aspects of orality and literacy'. In Lenker, U. and Meurman-Solin, A. (eds), *Connectives in History of English*. Amsterdam: John Benjamins Publishing, 289-309.

Kohonen, V. 1978. *On the Development of English Word Order in Religious Prose around 1000 and 1200 AD: a Quantitative Study of Word Order in Context*. (Publications of the Research Institute of the Åbo Akademi Foundation, 38.) Åbo Akademi: Åbo.

Koopman, W. 1990. 'Old English constructions with three verbs'. *Folia Linguistica Historica* 9.1 (1990), 271-300.

Koopman, W. 1992. 'Old English clitic pronouns: some remarks'. In Colman, F. (1992) (ed.), 44-87.

Koopman, W. F. 1985. 'Verb and particle combinations in Old and Middle English'. In Eaton, R., Fischer O., Koopman, W. and van der Leek, F. (eds), *Papers from the 4th International Conference on English Historical Linguistics: Amsterdam, 10-13 April 1985*. (Current Issues in Linguistic Theory, 41) John Benjamins, Amsterdam/Philadelphia. (1985), 109-21.

Koopman, W. F. 1998. 'Inversion after single and multiple topics in Old English'. In J. Fisiak and Krygier, M. (eds), *Advances in English historical linguistics*. (Trends in Linguistics, Studies and Monohraphs). Walter de Gruyter, 135-150.

Kroch, A and Taylor, A. 2000. *Penn-Helsinki Parsed Corpus of Middle English*, Second Edition (PPCME2). Available at http://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-2/index.html, accessed March 23, 2007.

Kroch, A. 2003. 'Language Change: Syntactic Change'. In Frawley, W. J. (ed.), *International Encyclopaedia of Linguistics* (second edition) vol.4. OUP, 397-401.

Kroch, A. 2008. 'Syntactic Change'. In Baltin, M. and Collins, Ch. (eds), *The Handbook of Contemporary Syntactic Theory*. Oxford: Wiley-Blackwell, 698-729.

Kroch, A., Taylor, A. and Ringe, D. 2000. 'The Middle English Verb-Second Constraint: A Case Study in Language Contact and Language Change. In. Herring, S. C, van Reenen, P., and Schøsler, L. (eds), *Textual Parameters in Older Languages*. John Benjamins Publishing, 353-391.

Krygier, M. 1998. 'The origin of Middle English *she* – An alternative hypothesis'. In Puppel, S. (ed.), *Scripta Manent*, Wydział Neofilologii UAM, 117-124.

Kuryłowicz, J. 1965. 'The evolution of grammatical categories'. *Diogenes* 51: 55–71. (Reprint: Kuryłowicz, 1975, *Esquisses linguisique* II, 38 – 54. München: W: Fink).

Labov, W. 1966. *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Laing, L. and Laing, J. 1979. *Anglo-Saxon England*. London: Routledge and Kegan Paul.

Laing, M. 2000. 'Never the twain shall meet: Early Middle English - the East-West divide'. In Taavitsainen, I., Nevalainen, T., Pahta, P. and Rissanen, M. (eds), *Placing Middle English in Context*. Berlin: Mouton de Gruyter, 97-124.

Laing, M. (ed.). 1989. *Middle English Dialectology: Essays on Some Principles and Problems*. Aberdeen: Aberdeen University Press.

Laker, S. 2002. 'An explanation for the changes kw-, hw- > χw- in the English dialects'. In Filppula, Klemola and Pitkänen (eds), 183–99.

Laker, S. 2009. 'An explanation for the early phonemicisation of a voice contrast in English fricatives'. *English Language and Linguistics* 13(2): 213–26.

Lakoff, R. 1972. 'Another look at drift'. In Stockwell, R.P. (ed.), *Linguistic change and generative theory: essays from the UCLA Conference on Historical Linguistics in the perspective of transformational theory, Febr. 1963*. Bloomington: India University Press, 172-198.

Lamers, M. J. A. and de Swart, P. 2012. 'The Interaction of Case, Word Order and Prominence: Language Production and Comprehension in a Cross-linguistic perspective'. In Lamers, M. and de Swart, P. (eds), *Case, Word Order and Prominence: Interacting Cues in Language Production and Comprehension*. Dordrecht: Springer, 1-16.

Lass, R. 1992. 'Phonology and Morphology'. In N. Blake (ed.), 23-155.

Lee, D. Y. W. 2001. 'Genres, register, text types, domains and styles: Clarifying the concepts and navigating a path through the BNC jungle'. *Language, Learning and Technology* Vol. 5, Num 3: 37-72. Available at http://llt.msu.edu/vol5num3/lee/default.html, accessed July 25, 2011.

Lehmann, W. P. 1973. *Historical Linguistics: An Introduction*. New York: Holt, Rinehart and Winston, Inc.

Lehmann, W. P. (ed.). 1978a *Syntactic typology: studies in the phenomenology of language*. University of Texas Press: Austin.

Lehmann, W.P. 1973. 'A structural principle of language and its implications'. *Language* 49: 47-66.

Lehmann, W.P. 1978b. 'The great underlying ground-plans'. In W.P. Lehmann (ed.), 3-55.

Li, Ch. N. (ed.). 1977. *Mechanisms of syntactic change*. Austin: University of Texas Press.

Lightfoot, D. 1979. 'Review of Li (1977)'. *Language* 55: 381-95.

Lightfoot, D. 1981. 'Explaining Syntactic Change'. In Hornstein, N. and Lightfoot, D. (eds), *Explanation in Linguistics: the Logical Problem of Language Acquisition*. London: Longman, 207-40.

Lightfoot, D. 1991. *How to set parameters: Arguments from language change*. Cambridge, Mass.: MIT Press.

Lightfoot, D. 1995. 'Why UG needs a learning theory: Triggering verb movement'. In Battye, A. and Roberts, I. (eds), 31-52.

Lightfoot, D. 1997. 'Shifting triggers and diachronic reanalysis'. In van Kemenade, A. and Vincent, N. (eds), 235-272.

Lightfoot, D. W. 1999. *The Development of Language: Acquisition, Change, and Evolution*. Oxford: Blackwell.

Lightfoot, D. W. (ed.). 2002. *Syntactic effects of morphological change*. Oxford: Oxford University Press.

Lightfoot, D. 2006a. 'Cuing a New Grammar'. In van Kemenade, A. and Los, B. (eds), 24-44.

Lightfoot, D. 2006b. *How New Languages Emerge*. Cambridge: Cambridge University Press.

Lightfoot, D. and Westergaard, M. 2007. 'Language acquisition and language change: Inter-relationships'. *Language and Linguistics Compass* 1/5: 396-415.

Lindkvist, H. 1912. *Middle-English Place-Names of Scandinavian Origin*. Upsala: The University Press.

Logan, D. 1983. *The Vikings in History*. Totowa: Barnes & Noble Books.

Loyn, H. 1994. *The Vikings in Britain*. Oxford: Blackwell Publishers.

Lutz, A. 2009. 'Celtic Influence on Old English and West Germanic'. *English Language and Linguistics* 13:2, 227-249.

Lutz, A. 2012. 'Language Contact in the Scandinavian Period'. In Nevalainen, T. and Traugott, E. (eds), *The Oxford handbook of the History of English.* Oxford: Oxford University Press, 508-517.

Lutz, A. 2013. 'Language Contact and Prestige'. *Anglia* 131 (4): 562-590.

Maiden M., Smith, J. Ch. and Ledgeway, A. (eds). 2011. *The Cambridge History of the Romance Languages*. Cambridge: Cambridge University Press.

Mallinson, G. and Blake, B. J. 1981. *Language typology: cross-linguistic studies in syntax.* North-Holland: Amsterdam.

Manfred, M. 1995. 'On the growing role of semantic and pragmatic features in Middle English'. In Fisiak, J. (ed.), *Linguistic Change under Contact Conditions*, Walter de Gruyter, 161-178.

Masayuki, O. 2010. 'On Stylistic Fronting in Middle English Prose'. In Kytö, M., Scahill, J., and Tanabe, H. (eds), *Language Change and Variation from Old English to Late Modern English: A Festschrift for Minoji Akimoto* (Linguistic Insights, Vol 114) Bern: Peter Lang AG, 103-116.

Matras, Y. 2009. *Language Contact*. Cambridge: Cambridge University Press.

Matras, Y. 2010. 'Contact Convergence and Typology'. In Hickey, R. (ed.), 66-85.

McEnery, T and Wilson, A. 1996. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

McEnery, T., Xiao, R. and Tono, Y. 2006. *Corpus-Based Language Studies: An Advanced resource book*. New York: Routledge.

McFadden, T. 2003. 'On Morphological Case and Word-Order Freedom'. In *Proceedings of the Twenty-Ninth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetic Sources of Phonological Patterns: Synchronic and Diachronic Explanations*, 295-306.

McFadden, T. 2005. 'OV-VO in English and the role of case-marking in word order'. *English Language and Linguistics* 9.1: 63-82.

McIntosh A., Samuels, M.L. and Benskin, M. 1986. *A Linguistic Atlas of Late Mediaeval English*. Aberdeen: Aberdeen University Press.

McMahon, A. and McMahon, R. 2003. 'Finding Families: Quantitative Methods in Language Classification'. *Transactions of the Philological Society* Volume 101, Issue 1, 7–55.

McMahon, A. 1994. *Understanding Language Change*. Cambridge: Cambridge University Press.

McWhorter, J. H. 2002. 'What happened to English?'. *Diachronica* 19: 217-272.

McWhorter, J. 2009. 'What else happened to English? A brief for the Celtic hypothesis'. *English Language and Linguistics*, 13: 163-191.

McWhorter, J. H. 2011. *Language Contact and Bilingualism: Linguistic Simplicity and Complexity: Why Do Languages Undress?* Berlin: Walter de Gruyter.

Meillet, A. 1912. *L'évolution des formes grammaticales*. Scientia (Rivista di scienza), Vol. XII (1912), n° XXVI, 6. (Reprint: Meillet, 1965, *Linguistique Historique et Linguistique Générale*, 130– 148. Paris: Librairie Honoré Champion, Editeur).

Miles, D. 2006. *The Tribes of Britain*. London: Phoenix.

Miller, D. G. 2002. *Nonfinite Structures in Theory and Change*. Oxford: Oxford University Press.

Miller, D. G. 2012. *External Influences on English: From its Beginnings to the Renaissance*. Oxford: Oxford University Press.

Millward, C. M. 1996. *A Biography of the English Language*. Harcourt: Harcourt Brace College.

Milroy, J. 1992a. *Linguistic Variation and Change: On the Historical Sociolinguistics of English*. Oxford/Cambridge, MA: Blackwell.

Milroy, J. 1992b. 'Middle English Dialectology'. In Blake, N. (ed.), 156-206.
Milroy, J. 1999. 'Toward a speaker-based account of language change'. In Jahr, E. (ed.), *Language Change: Advances in Historical Sociolinguistics*. Berlin: Mouton, 21-36.
Milroy, L. 1984. 'Comprehension and Context: Successful Communication and Communicative Breakdown'. In P. Trudgill (ed.), 7-31.
Milroy, L. 1987. *Language and Social Networks*, Second Edition. Language in Society 2. Oxford/ Cambridge, MA: Blackwell.
Mitchell, B. 1964. Syntax and Word Order in the Peterborough Chronicle, 1122-1154 in *Neuphilologische Mitteilungen* 65: 113-144.
Mitchell, B. and Robinson, F.C. 1983. *A Guide to Old English Revised with Texts and Glossary*. Oxford: Basil Blackwell Publisher Limited [1964].
Mitchell, B. 1985. *Old English Syntax*, 2 vols. Clarendon: Oxford.
Mitchell, B. 1994. 'The Englishness of Old English'. In Godden, M. R. Gray, D. and Hoad, T. (eds), *From Anglo-Saxon to Early Middle English: Studies Presented to E. G. Stanley*. Oxford: Clarendon, 163-81.
Mitchener, W. G. 2005. 'Simulating Language Change in the Presence of Non-Idealized Syntax'. In *Proceedings of the Second Workshop on Psychocomputational Models of Human Language Acquisition*. 10-19.
Mithun, M. 1992. 'Is Basic Word Order Universal?'. In Payne, D. L. (ed.), *Pragmatics of Word Order Flexibility* (Typological Studies in Language Vol 22). Amsterdam: John Benjamins Publishing Co., 15-61.
Mittendorf, I. and Poppe, E. 2000. 'Celtic contacts of the English progressive?'. In Tristram, H. L. C. (ed.), 117–45.
Molencki, R. 2011. 'New Prepositions and Subordinating Conjunctions of Romance Origin in Middle English'. In Fisiak, J. and Bator, M. (eds), 9-23.
Morris, Ch. D. 1981. 'Viking and Native in Northern England: A Case Study'. In Bekker-Nielsen, H., Foote, P. G. and Olsen, O. (eds), *Proceedings of the Eighth Viking Congress, Arhus, 24-31 August 1977* (Mediaeval Scandinavia Supplements 2). Odense: Odense University Press, 223-44.
Morris, Ch. D. 1982. 'The Vikings in the Birtish Isles'. In Farrell, R.T. (ed.), 70-94.
Mossé, F. 1938. *Histoire de la Forme Périphrastique Être + Participe Présent en Germanique. Deuxième Partie: Moyen-Anglais et Anglais Moderne.* Paris: Librairie C. Klincksieck.
Mugglestone, L. (ed.). 2006. *The Oxford History of English*. Oxford: Oxford University Press.
Neeleman, A. and Weerman, F. 1999. *Flexible syntax: A theory of case and arguments*. Dordrecht: Kluwer.
Newman, J. G. 2008. *The spread of the s-plural formative in Old and Middle English nouns* in Wełna, J. (ed.), *Warsaw studies in English Historical Linguistics*, Vol. 4 Institute of English Studies: University of Warsaw.
Newmeyer, F. J. 2003. "Formal and functional motivation for language change" in Hickey, R. (ed.), 18-36.
Nickel, G. 1966. *Die Expanded Form im Altenglischen: Vorkommen, Funktion und Herkunft der Umschreibung 'beon/wesan' + Partizip Präsens*. Neumünster: Karl Wachholz Verlag.
Noonan, M. 2010. 'Genetic Classification and Language Contact'. In Hickey, R. (ed.), 48-65.
O'Neil, W. 1978. 'The evolution of the Germanic inflectional systems: A study in the causes of languages change'. *Orbis* 27: 248-286.

Ogura, M. 2001. 'Perceptual factors and word order change in English'. *Folia Linguistica Historica* 22: 233-53.

Oppenheimer, S. 2006. *The Origins of the British*. London: Constable and Robinson Ltd.

Pagliuca, W. (ed.). 1994. *Perspectives on Grammaticalization* [Current Issues on Linguistic theory 109]. Amsterdam: John Benjamins.

Pintzuk, S. and Kroch, A. S. 1985. 'Reconciling an exceptional feature of Old English clause structure'. In Faarlund, J. T. (ed.), *Germanic Linguistics: Papers from a Symposium at the University of Chicago, April 24, 1985*. Indiana University Linguistics Club: Bloomington (1985), 87-111.

Pintzuk, S. and Kroch, A. S. 1989. 'The rightwards movement of complements and adjuncts in the Old English of Beowulf" (Revision of Pintzuk and Kroch (1985)'. *Language Variation and Change* 1 (1989), 115-43.

Pintzuk, S. and Taylor, A. 2006. 'The Loss of OV Order in the History of English'. In van Kemenade, A. and Los, B. (eds), 249-278.

Pinztuk, S. 1991. *Phrase structures in competition: Variation and Change in Old English word order.* Doctoral Dissertation, University of Pennsylvania.

Pinztuk, S. 1998. *Phrase Structures in Competition: Variation and Change in Old English Word Order* (Outstanding Dissertations in Linguistics). New York: Garland.

Platzack, C. 1995. 'The loss of verb second in English and French'. In Battye, A. and Roberts, I. (eds), 200-226.

Pollignton, S. 1989. *The warrior's way: England in the Viking age*. London: Blandford Press.

Poole, A. L. 1958. *Medieval England*. Oxford: Clarendon Press.

Poppe, E. 2003. 'Progress on the progressive? A report". In Tristram (ed.), 65–84.

Posner, R. 1997. *Linguistic change in French*. Oxford: Oxford University Press.

Poussa, P. 1982. 'The Evolution of Early Standard English: The Creolization Hypothesis'. *Studia Anglica Posnaniensia*, 14: 69-85.

Poussa, P. 1990. 'A contact-universals origin for periphrastic do, with special consideration of OE-Celtic contact'. In Adamson, S., Law, V., Vincent, N. and Wright, S. (eds), *Papers from the 5th International Conference on English Historical Linguistics, Cambridge, 6-9 April 1987*. Amsterdam: John Benjamins, 407-434.

Primus, B. 1998. 'The relative order of recipient and patient in the languages of Europe'. In Siewierska, A (ed.), *Constituent Order in the Languages of Europe*. Berlin: Mouton de Gruyter. 421-473.

Prins, A. A. 1948. 'French phrases in English'. *Neophilologus* 32, 1: 28-39; 2: 73-82.

Rankin, R. L. 2003. 'The Comparative Method (Methods for Studying Language Change'. In Joseph, B.D. and Janda, R. D. (eds), 183-212.

Renfrew, C. 1989. 'Models of change in language and archaeology'. *Transactions of the Philological Society* Volume 87:2, 102-155.

Richards, J. 2001. *Blood of the Vikings*. London: Hodder and Stoughton.

Richards, J. 2005. *Vikings: A Very Short Introduction*. Oxford: Oxford University Press.

Rissanen, M. 1992. 'The diachronic corpus as a window to the history of English'. In Svartvik, J. (ed.), *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 83, Stockholm, 4-8 August 1991*. Berlin: Mouton de Gruyter, 185-205.

Rissanen, M., Kytö, M., Kahlas-Tarkka, L., Kilpiö, M., Nevanlinna, S., Taavitsainen, I., Nevalainen, T., Raumolin-Brunberg, H. 1991. *The Helsinki Corpus of English Texts*. Department of Modern Languages, University of Helsinki.

Rissanen, M., Kytö, M. and Palander, M. (eds). 1993. *Early English in the Computer Age: Explorations through the Helsinki Corpus*. Berlin: Mouton de Gruyter.

Roberge, P. 2010. 'Contact and the History of Germanic Languages'. In Hickey, R. (ed.), 406-431.

Roberts, I. and Roussou, A. 2003. *Syntactic Change. A Minimalist Approach to Grammaticalization*. Cambridge: CUP.

Roberts, I. 1993. *Verbs and diachronic syntax: A comparative history of English and French*. Dordrecht: Kluwer.

Roberts, I. 1997. 'Directionality and word order change in the history of English'. In van Kemenade, A. and Vicent, N. (eds), 397-426.

Robinson, O. W. 1992. *Old English and Its Closest Relatives: A Survey of the Earliest Germanic Languages*. Stanford: Stanford University Press.

Roesdahl, E. 1991. *The Vikings*. London: The Penguin Press.

Romaine, S. 2002. 'Germanic Creoles'. In König, E. and van der Auwera, J. (eds), *The Germanic Language* [second edition]. Routledge, London and New York, 556-603.

Ross, M. 2003. 'Diagnosing prehistoric language contact'. In Hickey, R. (ed.), 174-198.

Rowland, T. H. 1994. *Anglo Saxon Northumbria*. Warkworth: Sandhill Press Ltd.

Ruiz Moneva, M. A. 1997. 'A Relevance Theory Approach to the Scandinavian Influence upon the Development of the English Language'. *Revista Alicantina de Estudios Ingleses* 10: 183-91.

Samuels, M. L. 1971. 'Kent and the Low Countries: some linguistic evidence'. In Aitken, A. J. McIntosh, A. and Pálsson, H. (eds), *Edinburgh Studies in English and Scots,* 3-19. London: Longman.

Samuels, M. L. 1985. 'The Great Scandinavian Belt'. In Eaton, R., Fischer, O., Koopman, W. and van der Leek, F. (eds), *Papers from the 4th International Conference on English Historical Linguistics*. Amsterdam: John Benjamins, 269-81. [Reprinted in Laing 1989: 106-15].

Santorini, B. and Kroch, A. 2007. 'The Verb-Second (V2) Phenomenon". *The Syntax of Natural Language: An Online Introduction Using Trees Program*. Available at http://www.ling.upenn.edu/~beatrice/syntax-textbook/, accessed May 5, 2015.

Sapir, E. 1921. *Language: an introduction to the study of speech*. New York: Harcourt, Brace & World.

Sawyer, P. H. 1971. *The Age of the Vikings* (Second Edition). London: Edward Arnold.

Schiffman, H. F. 2010. 'Language Contact in South Asia'. In Hickey, R. (ed.), 738-756.

Schmidt, J. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen.* Weimar: H. Böhlau.

Schøsler, L. 2001. 'From Latin to Modern French: Actualization and Markedness'. In Andersen, H. (ed.), *Actualization: Linguistic Change in Progress*. Amsterdam: John Benjamins Publishing Company, 169-186.

Schrijver, P. 2002. 'The rise and fall of British Latin: evidence from English and Brittonic'. In Filppula et al. (eds), 88–110.

Schumacher, S. 2007. 'Die Deutschen und die Nachbarstämme: Lexikalische und strukturelle Sprachkontaktphänomene entlang der keltisch-germanischen Übergangszone'. *Keltische Forschungen* 2: 167-207.

Scragg, D. G. 2008. 'The Vercelli Homilies and Kent'. In Blanton, V. and Scheck, H. (eds), *Intertexts: Studies in Anglo-Saxon Culture Presented to Paul E. Szarmach* (Medieval and Renaissance Texts and Studies, Volume 334). Tempe: ACMRS, 369-380.

Scragg, D. G. (ed.). 1992. *The Vercelli Homilies and Related Texts*. Oxford: Oxford University Press.

Seoane, E. 2006. 'Information Structure and Word Order Change: The Passive as an Information-rearranging Strategy in the History of English'. In van Kemenade, A. and Los, B. (eds), 360-391.

Serjeantson, M. S. 1935. *History of Foreign Words in English*. London: Routledge and Kegan Paul.

Siewierska, A. 1988. *Word order rules*. Croom Helm: London.

Siewierska, A. and Bakker, D. 1996. 'The distribution of subject and object agreement and word order type'. *SiL* 20: 115-61.

Siewierska, A. and Bakker, D. 2008. 'Case and alternative strategies; word order and agreement marking'. In Malchukov, A. and Spencer, A. (eds), *The Oxford Handbook of Case*. New York: OUP, 290-303.

Smail, R. C. 1958. 'Art of war'. In Poole, A. L. (ed.), 128-167.

Smith, C. A. 1893. 'The order of words in Anglo-Saxon prose'. *Publications of the Modern Language Association of America* 8 (1983), 210-44.

Song, J. J. 2001. *Linguistic Typology: Morphology and Syntax*. Harlow: Pearson Education Limited.

Song, J. J. 2011b. 'Word Order Typology'. In Song, J. J. (ed.), 253-279.

Song, J. J. (ed.). 2011a. *Handbook of Linguistic Typology*. Oxford: Oxford University Press.

Stafford, P. 2003. 'Kings, Kingships, and Kingdoms'. In Davies, W. (ed.), 11-39.

Stockwell, J. and Minkova, D. 2000. 'What happened to Old English clitic pronouns and why?'. In Dalton-Puffer, Ch. and Ritt, N. (eds), *Words, Structure, Meaning, Function*. Berlin: Mouton de Gruyter, 289-307.

Stockwell, R. P. and Minkova, D. 1990. 'Verb Phrase conjunction in Old English'. In Andersen, H and Koerner, K. (eds), *Linguistics 1987: Papers from the 8th International Conference in Historical Linguistics (8. ICHL) (Lille, 31 August – 4 September 1987)* (Current Issues in Linguistic Theory, 66.) John Benjamins: Amsterdam/Philadelphia (1990), 499-515.

Stockwell, R. P. and Minkova, D. 1991. 'Subordination and word order change in the history of English'. In Kastovsky, D. (ed.), *Historical English Syntax*. (Topics in English Linguistics, 2.) Mouton de Gruyter: Berlin/New York (1991), 367-408.

Stockwell, R. P. and Minkova, D. 1992. 'Poetic influence on prose word order in Old English'. In Colman, F. (1992) (ed.), 142-54.

Stockwell, R. P. 1977. 'Motivations for Exbraciation in Old English'. In Li, Ch. N. (ed.), 291-314.

Stockwell, R. P. 1981. 'On the history of the verb-second rule in English'. In Fisiak, J. (ed.), 575-592.

Story, J. 1995. *Charlemagne and Northumbria: the influence of Francia on Northumbrian politics in the late eighth and the early ninth centuries*. Durham theses, Durham University. Available at http://etheses.dur.ac.uk/1460/, accessed Oct 27, 2013.

Strang, B. M. H. 1970. *A History of English*. Methuen: London.

Strohm, P. (ed.). 2007. *Oxford Twenty-first century approaches to Literature: Middle English*. Oxford: Oxford University Press.

Stubbs, M. 1996. *Text and corpus analysis: Computer assisted studies of language and culture*. Oxford: Blackwell.

Swales, J. 1990. *Genre Analysis: English in Academic and Research Settings*. New York: Cambridge: Cambridge University Press.

Swan, T. 1991. 'Adverbial shifts: Evidence from Norwegian and English'. In Kastovsky, D. (ed.), 409-438.

Swan, T. 1994. 'A note on Old English and Old Norse initial adverbials and word-order with special reference to sentence adverbials'. In Swan et al. (eds), 233-270.

Swan, T., Mørck, E. and Jansen, O. (eds). 1994. *Language Change and language structure: older Germanic languages in a comparative perspective*. Walter de Gruyter.

Sykes, B. 2007. *Blood of the Isles: Exploring the genetic roots of our tribal history*. London: Transworld Corgi.

Szmrecsanyi, B. 2013. 'The great regression: Genitive variability in Late Modern English news texts'. In Börjars, K., Denison, D. and Scott, A.K. (eds), *Morphosyntactic Categories and the Expression of Possession.* Amsterdam: Benjamins, 59-88.

Taylor, A., Warner, A., S. Pintzuk, and Beths, F. 2003. *The York-Toronto-Helsinki Parsed Corpus of Old English Prose* (YCOE). (http://www-users.york.ac.uk/~lang22/YCOE/YcoeHome.htm, accessed April 2, 2007)

Thomason, S. and Kaufman, T. 1988. *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.

Thomason, S. 2001. *Language Contact*. Edinburgh: Edinburgh University Press Ltd.

Thomason, S. 2003. 'Contact as a Source of Language Change'. In Joseph, B.D. and Janda, R. D. (eds), 687-712.

Thomason, S. G. and Everett, D. L. 2005. 'Pronoun Borrowing'. *Proceedings of the Berkeley Linguistics Society* 27: 305-15.

Thorson, P. 1936. *Anglo-Norse Studies: An Inquiry into the Scandinavian Elements in the Modern English Dialects, Part I*. Amsterdam: N.V. Swets en Zeitlinger.

Tognini-Bonelli, E. 2001. *Corpus Linguistics at Work*. Studies in Corpus Linguistics. Amsterdam: John Benjamins B.V.

Townend, M. 2002. *Language and history in Viking Age England: linguistic relations between speakers of Old Norse and Old English*. Turnhout: Brepols.

Townend, M. 2006. 'Contacts and Conflicts: Latin, Norse, and French'. In Mugglestone, L. (ed.), 61-85.

Traugott, E. 1992. 'Syntax'. In Hogg, R. (ed.), 168-289.

Traugott, E. C. 1972. *The History of English Syntax: A Transformational Approach to the History of English Sentence Structure*. New York: Holt, Rinehart and Winston, Inc.

Trips, C. 2002. *From OV to VO in Early Middle English*. Amsterdam: John Benjamins Publishing Company.

Tristram H. 2007. 'Why Don't the English Speak Welsh?'. In Higham, N. J. (ed.), *Britons in Anglo-Saxon England*. Woodbridge: Boydell Press, 192-214.

Tristram, H. 1999. 'The Celtic Englishes' – Zwei grammatische Beispiele zu Problem des Sprachkontsktes zwischen dem Englischen und den keltischen Sprachen'. In Zimmer, S., Kodderitzsch, R. and Wigger, A. (eds), *Akten des zweiten Deutschen Keltologen-Symposiums* (Bonn, 2-4 April 1997). Tubingen: Max Niemeyer Verlag, 254-276.

Tristram, H. L. C. (ed.). 2000. *Celtic Englishes II*. Heidelberg: Universitätsverlag C. Winter.

Tristram, H. L. C. (ed.). 2003. *Celtic Englishes III*. Heidelberg: Universitätsverlag C. Winter.

Tristram, H. L. C. 2004. 'Diglossia in Anglo-Saxon England, or what was spoken Old English like?'. In J. Fisiak (ed.), *Studia Anglica Posnaniensia* 40, 87–110.

Tristram, H. 2002. 'Attrition of Inflections in English and in Welsh'. In Filppula, M. et al. (eds), 111-149.

Trotter, D. 2012. 'English in Contact: Middle English creolization'. In Bergs, A. and Brinton, L. *Historical Linguistics of English* (HSK 34.1) New York and Berlin: de Gruyter, 1781-1793.

Trudgill, P. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, P. 1984. *Applied Sociolinguistics*. London: Academic Press.

Trudgill, P. 1986. *Dialect in Contact*. Oxford: Basil Blackwell.

Trudgill, P. 2002. 'Third-person singular zero: African-American English, Anglian dialects and Spanish persecution in the Low Countries'. In Fisiak, J. and Trudgill, P. (eds), *East Anglian English*, Cambridge: D. S. Brewer, 179-186. (First published in 2001)

Trudgill, P. 2010. *Investigations in Sociohistorical Linguistics: Stories of Colonisation and Contact*. Cambridge: Cambridge University Press.

Trudgill, P. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.

Upton, C. 2006. 'Modern Regional English in the British Isles'. In Mugglestone, L. (ed.), 305-333.

Vennemann, T. 1974. 'Topics, subjects and word order: from SXV to SVX via TVX'. In Anderson, J. M. and Jones, C. (eds), *Historical Linguistics I: Syntax, Morphology, Internal and Comparative Reconstruction: Proceedings of the First International Conference on Historical Linguistics, Edinburgh 2nd-7th September 1973*. (North-Holland Linguistic Series, 12a) North-Holland, Amsterdam/Oxford/American Elsevier, New York (1974), 339-76.

Vennemann, T. 1984. 'Verb second, verb late, and the brace construction'. In Fisiak, J. (ed.), 627-36.

Vennemann, T. 2000. 'English as a "Celtic" language: Atlantic influences from above and from below'. In Tristram, H. L. C. (ed.), 399–406.

Vennemann, T. 2001. 'Atlantic Semitica: Structural contact features in Celtic and English'. In Brinton, L.J. (ed.), *Historical Linguistics 1999: Selected Papers from the 14th International Conference on Historical Linguistics, Vancouver, 9-13 August 1999*. Amsterdam and New York: Benjamins, 351-369.

Vennemann, T. 2011. 'English as a contact language: Typology and comparison'. *Anglia* 129: 217-257.

Vezzosi, L. 2012. 'English in Contact: Latin'. In Bergs, A. and Brinton, L. J. (eds), *English Historical linguistics: An International Handbook, Vol II* (Handbooks of Linguistics and Communication Science 34.2) Berlin and New York: De Gruyter, 1703-1719.

Visser, F. Th. 1963. *An Historical Syntax of the English Language*, Vol.1. Leiden: E. J. Brill.

Wainwright, E. T. 1975. *Scandinavian England*. Chichester: Phillimore and Co. Ltd.

Wakelin, M. 1988. *The Archaeology of English*. London: B.T. Batsford Ltd.

Wakelin, M. F. 1977. *English Dialects: An Introduction*. Rev. Edn. London: The Athlone Press.

Wales, K. 2006. *Northern English: A Cultural and Social History*. Cambridge: Cambridge University Press.

Weerman, F. 1989. *The verb-second conspiracy: A synchronic and a diachronic analysis*. Dordrecht: Foris.

Weerman, F. 1993. 'The diachronic consequences of first and second language acquisition: The change from OV to VO'. *Linguistics* 31,5 (327): 903-932.

Weinreich, U. 1968. *Languages in contact.* The Hague: Mouton. [Originally published in 1953 by the Linguistic Circle of New York]

Weinstock, H. 2006. 'Romanismen im Englischen/ Romanismes en anglais'. In Ernst, G., Gleßgen, M-D., Schmitt, Ch. and Schweickard, W. (eds), *Romanische Sprachgechichte/ Histoire linguistique de la Romania* 3 vols, Berlin: Walter de Gruyter, 1834-41.

White, D. L. 2002. 'Explaining innovations of Middle English: What, where, and why'. In Filppula, M. et al. (eds), 153-74.

White, D. L. 2003. 'Brittonic influence in the reductions of Middle English nominal morphology'. In Tristram, H. L. C. (ed.), 29-45.

Whitelock, D. 1974. *The Beginnings of English Society.* Harmondsworth: Penguin Books Ltd.

Whitelock, D., Douglas, D. C., Lemmon, C. H. and Burlow, F. 1966. *The Norman Conquest. Its Setting and Impact.* London: Eyre & Spottiswoode.

Whitney, W. D. 1979. *The Life and Growth of Language: An Outline of Linguistic Science*. New York: Dover. [Originally published in 1875.]

Wiemer, B. and Bisang, W. 2004. 'What makes grammaticalization? An appraisal of its components and its fringes'. In W. Bisang, N. Himmelmann, B. Wiemer, *What Makes Grammaticalization?* Berlin: Walter de Gruyter, 3-20.

Williams, A., Smyth, A. P. and Kirby, D. P. 1991. *A biographical dictionary of Dark Age Britain: England, Scotland and Wales c. 500-c.1050.* London: B. A. Seaby Ltd.

Williams, J. M. 1975. *Origins of the English Language: A Social and Linguistic History.* New York: The Free Press.

Wilson, D. M. 1976. 'The Scandinavians in England'. In Wilson, D. M. (ed.), *The Archaeology of Anglo-Saxon England.* Cambridge: Cambridge University Press, 393-403.

Winter-Froemel, E. 2008. 'Towards a comprehensive view of language change: Three recent evolutionary approaches'. In Detges, U. and Waltereit, R. (eds), *The Paradox of Grammatical Change: Perspectives from Romance.* Amsterdam: John Benjamins Publishing, 215-250.

Wolfram, W. and Schilling-Estes, N. 1998. *American English: Dialects and Variation*. Oxford: Blackwell Publishers.

Wolfram, W. and Shilling-Estes, N. 2003. 'Dialectology and Linguistic Diffusion'. In Joseph, B.D. and Janda, R. D. (eds), 713-35.

Workman, S. K. 1972. *Fifeenth century translations as an influence on English prose.* New York: Octagon Books.

Wormald, C. P. 1982. 'Viking Studies: Whence and Whither?'. In Farrell, R.T. (ed.), 128-153.

Wurff, W. van der 1990. *Diffusion and Reanalysis in Syntax*. Dissertation, University of Amsterdam.

Wurzel, W.U. 1989. *Inflectional Morphology and Naturalness*. Reidel, Dordrecht, Netherlands. / Dordrecht: Kluwer./ Dordrecht: Reidel

Yamane, T. 1973. *Statistics: An introductory analysis*. New York: Harper & Row.

Appendices

APPENDIX 1A

Text information for YCOE

code	date	dialect	Title	details (edition)	genre	ADDITIONAL INFO/TRANSLATION
coalex.o23	990-1010	Anglian NM	Alexander's Letter to Aristotle	Orchard, Andrew P.M. 1995. *Pride and Prodigies: Studies in the Monsters of the "Beowulf" Manuscript*. Pp. 224-52. Cambridge: D.S. Brewer.	Travelogue	Translation from Latin
cobede.o2	1050-1099	Anglian NM	Bede's History of the English Church	Miller, Thomas. 1959-1963 (1890-1898). *The Old English Version of "Bede's Ecclesiastical History of the English People"*. EETS 95, 96, 110, 111. London: OUP.	History	Translation from Latin
coblick.o23	990-1010	Anglian NM	Blickling Homilies	Morris, Richard. 1967 (1874-1880). *The Blickling Homilies*. EETS 58, 63, 73. London: Trübner.	Homilies, biographies/lives	
codocu3.o3_1486-7 (originally codocu3.o3)	975-1016	Anglian NM	Charters and Wills	Whitelock, Dorothy. 1930. *Anglo-Saxon Wills*. Cambridge: CUP.	Charters and wills	Clipped from a bigger chunk. No information on translation
coherbar	1000-1050	Anglian NM	Herbarium	de Vriend, Hubert Jan. 1984. *The Old English Herbarium and Medicina de quadrupedibus*. EETS 286:30-233. London: OUP.	Handbooks, medicine	Translation from Latin
colancu.o23	990-1050	Anglian NM	Lacnunga	Grattan, John Henry Grafton and Charles Singer, eds. 1952. *Anglo-Saxon Magic and Medicine*. Publications of the Wellcome Historical Medical Museum n.s. 3. London: OUP.	Handbooks, medicine	Translation from Latin
colaece.o2	940-960	Anglian NM	Bald's Leechbook	Cockayne, Oswald. 1864-1866. *Leechdoms, Wortcunning and Starcraft of Early England*. Rolls Series 35, vol. 1. 70-324. London: Her Majesty's Stationery Office. Reprinted Wiesbaden, Germany: Kraus Reprint Ltd. 1965.	Handbooks, medicine	Translation from Latin

comarvel.o23	990-1010	Anglian NM	Marvels of the East	Orchard, Andrew P.M. 1995. *Pride and Prodigies: Studies in the Monsters of the"Beowulf" Manuscript.* Pp. 184-202. Cambridge: D.S. Brewer.	Geography	Translation from Latin
coquadru.o23	1000-1050	Anglian NM	Quadrupedibus	de Vriend, Hubert Jan. 1984. *The Old English Herbarium and Medicina de quadrupedibus*. EETS 286: 234-73. London: OUP.	Handbooks, medicine	Translation from Latin
cochad.o24	1100-1150	Anglian AM	Other Saints' Lives, The Life of Saint Chad	Vleeskruyer, Rudolf. 1953. *The Life of Saint Chad: An Old English Homily*. Amsterdam: North-Holland.	Biography, lives	Translation from Latin
codocu2.o2_M (originally codocu2.o2)	899-904	Anglian AM	Charters and Wills	Robertson, A.J. 1956 (1939). *Anglo-Saxon Charters*. Cambridge: CUP.	Charters and wills	Clipped from a bigger chunk. No information on translation
codocu3.o23	743-883	Anglian AM	Charters and Wills	(1) Harmer, F.E. 1914. *Select Historical Documents of the Ninth and Tenth Centuries*. Cambridge: CUP. (2) Robertson, A.J. 1956 (1939). *Anglo-Saxon Charters*. Cambridge: CUP.	Charters and wills	Excluded for time normalisation. No information on translation
cogregdC.o24	1050-1099	Anglian AM	Gregory's Dialogues	Hecht, Hans. 1965 (1900-1907). *Bischof Wærferth von Worcester Übersetzung der Dialoge Gregors des Grossen.* Bibliothek der Angelsaechsischen Prosa, V. Darmstadt: Wissenschaftliche Buchgesellschaft.	Biography and lives	Translation from Latin
comart3.o23	990-1010	Anglian AM	Martyrology	Kotzor, G. 1981. *Das Alternglische Martyrologium, vol. II.* Bayerische Akademie der Wissenschaften, Philosophisch-Historische Klasse. Abhandlunge, Neue Folge, Heft 88/2. München: Verlag der Bayerischen Akademie der Wissenschaften.	Biography and lives	
codocu1.o1_K (originally codocu1.o1)	805-850	Kentish K	Charters and Wills	(1) Harmer, F.E. 1914. *Select English Historical Documents of the Ninth and Tenth Centuries*. Cambridge: CUP. (2) Robertson, A.J. 1956 (1939). *Anglo-Saxon Charters*. Cambridge: CUP.	Charters and wills	Excluded for time normalisation; Clipped from a bigger chunk. No information on translation
codocu2.o2_K (originally codocu 2.o2)	843-924	Kentish K	Charters and Wills	Harmer, F.E. 1914. *Select Historical Documents of the Ninth and Tenth Centuries.* Cambridge: CUP.	Charters and wills	Excluded for time normalisation; Clipped from a bigger chunk. No information on translation

codocu3.o3_1506 (originally codocu3.o3)	941-958	Kentish K	Charters and Wills	Robertson, A.J. 1956 (1939). *Anglo-Saxon Charters*. Cambridge: CUP.	Charters and wills	Clipped from a bigger chunk. No information on translation
coeluc1	1140-1160	Kentish K	Honorius of Autun, Elucidarium	Warner, Rubie D.-N. 1917 (1971). *Early English Homilies from the 12th Century Ms. Vespasian D.XIV.* EETS 152: 140-3. London: Trübner.	Religious treatise	This text is included in the PPCME2 as part of the Kentish Homilies (cmkentho.m1) Translation from Latin
coadrian.o34	840-860	West Saxon WS	Adrian and Ritheus	Cross, James E. and Thomas D. Hill. 1982. *The "Prose Solomon and Saturn" and "Adrian and Ritheus"*. Pp. 35-40. Toronto, Buffalo, London: University of Toronto Press.	Religious treatise	Excluded for time normalisation
coaelhom.o3	no date	West Saxon WS	Ælfric's Homilies Supplemental	Pope, J.C. 1968. *Homilies of Ælfric, A supplementary Collection*. Early English Text Society, 260. London: OUP.	Homilies	Excluded for time normalisation
coaelive.o3	1000-1015	West Saxon WS	Ælfric's Lives of Saints	Skeat, Walter William. 1966 (1881-1900). *Ælfric's Lives of Saints*. EETS 76, 82, 94, 114. London: OUP.	Biography, lives	
coappolo.o3	1040-1060	West Saxon WS	Apollonius of Tyre	Goolden, Peter. 1958. *The Old English "Apollonius of Tyre"*. London: OUP.	Fiction	Translation from Latin
cobenrul.o3	1000-1050	West Saxon WS	Benedictine Rule	Schröer, Arnold. 1885-1888. *Die angelsæchsischen Prosabearbeitungen der Benediktinerregel.* Bibliothek der Angelsæchsischen Prosa, II. Kassel. Reprinted with appendix by H. Gneuss (Darmstadt 1964).	Rule	Translation from Latin
coboeth.o2	940-960	West Saxon WS	Boethius, Consolation of Philosophy	Sedgefield, Walter John. 1899. *King Alfred's Old English Version of Boethius de Consolatione Philosophiae*. Oxford: Clarendon Press. Reprinted Darmstadt 1968.	Philosophy	Body of the text is a translation from Latin
cobyrhtf.o3	1040-1060	West Saxon WS	Byrhtferth's Manual	Baker, Peter S. and Michael Lapidge 1995. *Byrhtferth's Enchiridion.* EETS s.s. 15. Oxford: OUP.	Science	Translation from Latin
cocanedgX	1050-1075	West Saxon WS	Canons of Edgar	Fowler, Roger. 1972. *Wulfstan's Canons of Edgar*. EETS 266. London: OUP.	Ecclesiastical laws	
cocathom1.o3	990-1010	West Saxon WS	Ælfric's Catholic Homilies I	Clemoes, P. 1997. *Ælfric's Catholic Homilies: The First Series*. EETS s.s. 17. Oxford: OUP.	Homilies	
cocathom2.o3	990-1010	West Saxon WS	Ælfric's Catholic Homilies II	Godden, M. 1979. *Ælfric's Catholic Homilies: The Second Series*. EETS s.s. 5. London: OUP.	Homilies	

cochronA.o23	890-999	West Saxon WS	Anglo-Saxon Chronicle A	Plummer, Charles. 1965 (1892-1899). *Two of the Saxon Chronicles Parallel*. Oxford: Clarendon Press. Reissued D. Whitelock, Oxford 1952.	History	Excluded for time normalisation; Where A indicates one of the many scribes
cocura.o2	885-899	West Saxon WS	Cura Pastoralis	Sweet, Henry. 1958 (1871). *King Alfred's West-Saxon Version of Gregory's Pastoral Care.* EETS 45, 50. London: OUP.	Religious treatise	Excluded for time normalisation; Defective section 33 replaced by Cotton Tiberius B.XI, see cocuraC. Translation from Latin
coepigen.o3	1050-1099	West Saxon WS	Ælfric's Epilogue to Genesis	Crawford, Samuel J. 1922. *The Old English Version of the Heptateuch. Ælfric's Treatise on the Old and New Testament and His Preface to Genesis*. EETS 160: 333-76. London: OUP.	Epilogue	
coeuphr	1000-1015	West Saxon WS	Saint Euphrosyne	Skeat, Walter William. 1966 (1881-1900). *Ælfric's Lives of Saints*. EETS 76, 82, 94, 114: 334-54. London: OUP.	Biography and lives	Not by Ælfric according to Clemoes
coeust	1000-1015	West Saxon WS	Saint Eustace and his Companions	Skeat, Walter William. 1966 (1881-1900). *Ælfric's Lives of Saints*. EETS 76, 82, 94, 114: 190-218. London: OUP.	Biography and lives	Not by Ælfric according to Clemoes
coexodusP	1060-1090	West Saxon WS	Exodus	Crawford, Samuel J. 1922. *The Old English Version of the Heptateuch. Ælfric's Treatise on the Old and New Testament and His Preface to Genesis.* EETS 160: 458-60. London: OUP. Reprinted with additions by N.R. Ker 1969.	Bible	No information on translation
cogenesisC	1040-1060	West Saxon WS	Genesis	Crawford, Samuel J. 1922. *The Old English Version of the Heptateuch. Ælfric's Treatise on the Old and New Testament and His Preface to Genesis.* EETS 160: 444-56. London: OUP. Reprinted with additions by N.R. Ker 1969.	Bible	No information on translation
cogregdH.o23	1000-1050	West Saxon WS	Gregory's Dialogues	Hecht, Hans. 1965 (1900-1907). *Bischof Wærferth von Worcester Übersetzung der Dialoge Gregors des Grossen.* Bibliothek der Angelsaechsischen Prosa, V. Darmstadt: Wissenschaftliche Buchgesellschaft.	Biography, lives	Translation from Latin
coinspolD.o34	1000-1015	West Saxon WS	Institutes of Polity	Jost, K. 1959. "Die 'Institutes of Polity, Civil and Ecclesiastical.'" *Swiss Studies in English* 47. Bern.	Ecclesiastical laws	
colawaf.o2	890-999	West Saxon WS	Laws of Alfred	Liebermann, F. 1903-16. Die Gesetze der Angelsachsen. Halle. Reprinted Aalen 1960.	Laws	Excluded for time normalisation.

colawafint.o2	890-999	West Saxon WS	Alfred's Introduction to Laws	Liebermann, F. 1903-16. Die Gesetze der Angelsachsen. Halle. Reprinted Aalen 1960.	Laws	Excluded for time normalisation.
colawine.ox2	1090-1110	West Saxon WS	Laws of Ine	Liebermann, F. 1903-16. Die Gesetze der Angelsachsen. Halle. Reprinted Aalen 1960	Laws	
colsigef.o3	1140-1160	West Saxon WS	Ælfric's Letter to Sigefyrth	Assmann, Bruno. 1889. *Angelsaechsische Homilien und Heiligenleben*. Bibliothek der Angelsaechsischen Prosa, III. Kassel: Wissenschaftliche Buchgesellschaft. Reprinted with an introduction by P. Clemoes, Darmstadt 1964.	Religious treatise	Translation from Latin
colsigewB	1150-1199	West Saxon WS	Ælfric's Letter to Sigeweard (B)	Crawford, Samuel J. 1922. *The Old English Version of the Heptateuch. Ælfric's Treatise on the Old and New Testament and His Preface to Genesis*. EETS 160: 18-33, 39-51. London: OUP. Reprinted with additions by N.R. Ker 1969.	Religious treatise	No information on translation
colwgeat	1050-1160	West Saxon WS	Ælfric's Letter to Wulfgeat	Assmann, Bruno. 1889. *Angelsaechsische Homilien und Heiligenleben*. Bibliothek der Angelsaechsischen Prosa, III. Kassel: Wissenschaftliche Buchgesellschaft. Reprinted with an introduction by P. Clemoes, Darmstadt 1964.	Religious treatise	Translation from Latin
colwsigeT	1050-1075	West Saxon WS	Ælfric's Letter to Wulfsige	Fehr, B. 1914. *Die Hirtenbriefe Aelfrics in Altenglischer und Lateinischer Fassung*. Bibliothek der Angelsaechsischen Prosa, IX: 1-34. Hamburg: Verlag von Henri Grand. Reprinted with a supplement by P. Clemoes, Darmstadt 1966.	Religious treatise	No information on translation
colwsigeXa.o34	1040-1099	West Saxon WS	Ælfric's Letter to Wulfsige	Fehr, B. 1914. *Die Hirtenbriefe Aelfrics in Altenglischer und Lateinischer Fassung*. Bibliothek der Angelsaechsischen Prosa, IX: 1-34. Hamburg: Verlag von Henri Grand. Reprinted with a supplement by P. Clemoes, Darmstadt 1966.	Religious treatise	
colwstan1.o3	1040-1099	West Saxon WS	Ælfric's First Letter to Wulfstan	Fehr, B. 1914. *Die Hirtenbriefe Aelfrics in Altenglischer und Lateinischer Fassung*. Bibliothek der Angelsaechsischen Prosa, IX: 68-145. Hamburg: Verlag von Henri Grand. Reprinted with a supplement by P. Clemoes, Darmstadt 1966.	Religious treatise	Translation from Latin
colwstan2.o3	1040-1099	West Saxon WS	Ælfric's First Letter to Wulfstan	Fehr, B. 1914. *Die Hirtenbriefe Aelfrics in Altenglischer und Lateinischer Fassung*. Bibliothek der Angelsaechsischen Prosa, IX: 68-145. Hamburg: Verlag von Henri Grand. Reprinted with a supplement by P. Clemoes, Darmstadt 1966.	Religious treatise	Translation from Latin

comary	1000-1015	West Saxon WS	Mary of Egypt	Skeat, Walter William. 1966 (1881-1900). *Ælfric's Lives of Saints*. EETS 76, 82, 94, 114: 2-52. London: OUP.	Biography, lives	Not by Ælfric according to Clemoes
cootest.o3	1000-1099	West Saxon WS	Heptateuch	Crawford, Samuel J. 1922. *The Old English Version of the Heptateuch. Ælfric's Treatise on the Old and New Testament and His Preface to Genesis*. EETS 160. London: OUP. Reprinted with additions by N.R. Ker 1969.	Bible	Translation from Latin
coprefcath1.o3	990-1010	West Saxon WS	Ælfric's Preface to Catholic Homilies I	Clemoes, P. 1997. *Ælfric's Catholic Homilies: The First Series*. EETS s.s. 17: 174-77. Oxford: OUP.	Preface	
coprefcath2.o3	990-1010	West Saxon WS	Ælfric's Preface to Catholic Homilies II	Godden, M. 1979. *Ælfric's Catholic Homilies: The Second Series*. EETS s.s. 5: 1-2. London: OUP.	Preface	
coprefcura.o2	885-899	West Saxon WS	Preface to Cura Pastoralis	Sweet, Henry. 1958 (1871). *King Alfred's West-Saxon Version of Gregory's Pastoral Care*. EETS 45, 50: 3-9. London: OUP.	Preface	Excluded for time normalisation.
coprefgen.o3	1050-1099	West Saxon WS	Ælfric's Preface to Genesis	Crawford, Samuel J. 1922. *The Old English Version of the Heptateuch. Ælfric's Treatise on the Old and New Testament and His Preface to Genesis*. EETS 160: 76-80. London: OUP. Reprinted with additions by N.R. Ker 1969.	Preface	
copreflives.o3	1000-1015	West Saxon WS	Ælfric's Preface to Lives of Saints	Skeat, Walter William. 1966 (1881-1900). *Ælfric's Lives of Saints*. EETS 76, 82, 94, 114: 4-6. London: OUP.	Preface	
cosevensl	1000-1015	West Saxon WS	The Seven Sleepers	Magennis, Hugh 1994. *The Anonymous Old English Legend of the Seven Sleepers*. Durham Medieval Texts 7. Durham.	Biography, lives	
cosolilo	1140-1160	West Saxon WS	St Augustine's Soliloquies	Endter, W. 1922. *König Alfreds des Grossen Bearbeitung der Soliloquien des Augustinus*. Bibliothek der Angelsaechsischen Prosa, 11. Darmstadt: Wissenschaftliche Buchgesellschaft. Reprinted Darmstadt 1964. Corrections by Carnicelli, T.A. 1969. *King Alfred's Version of St. Augustine's Soliloquies*. Cambridge, MA: Harvard University Press.	Religious treatise	Translation from Latin
cotempo.o3	990-1010	West Saxon WS	De Temporibus Anni	Henel, Heinrich. 1970 (1942). *Ælfric's De Temporibus Anni*. EETS 213. London: OUP.	Science, astronomy	Translation from Latin
cowsgosp.o3	1000-1050	West Saxon WS	West-Saxon Gospels	Skeat, Walter William. 1871-1887. *The Four Gospels in Anglo-Saxon, Northumbrian and Old Mercian Versions*. Cambridge: CUP. Reprinted Darmstadt 1970.	Bible	Translation from Latin
cowulf.o34	no date	West Saxon WS	The Homilies of Wulfstan	Bethurum, Dorothy. 1957. *The Homilies of Wulfstan*. Oxford: Clarendon Press.	Homilies	Excluded for time normalisation.

APPENDIX 1B

Text information for PPCME2

code	date	dialect	title	details (edition)	genre	ADDITIONAL INFO	TRANSLATION
cmrolltr.m24	1250-1350	North	Richard Rolle, Prose Treatises from the Thornton Ms.	Perry, George G. 1921. *English prose treatises of Richard Rolle de Hampole.* EETS O.S. 20. London: Oxford University Press.	Religious treatise		Translation from Latin
cmrollep.m24	1250-1350	North	Richard Rolle, Epistles (Ego Dormio, The Commandment, The Form of Living)	Allen, Hope E. 1931. *English writings of Richard Rolle, hermit of Hampole.* Oxford: Clarendon.	Religious treatise		?
cmbenrul.m3	1350-1420	North	The Northern Prose Rule of St. Benet	Kock, Ernst A. 1902. *The Northern prose version of the Rule of St. Benet.* In Ernst A. Kock (ed.), *Three Middle-English versions of the Rule of St. Benet and two contemporary rituals for the ordination of nuns*. EETS O.S. 120. London: K. Paul, Trench, Trübner & Co	Rule	This manuscript represents the earliest extensive Northern prose. According to Kock, it is not a close translation of the Latin rule.	
cmedthor.m34	1350-1420	North	The Mirror of St. Edmund (Thornton Ms.)	Perry, George G. 1969. *The mirror of St. Edmund.* In George G. Perry (ed.), *Religious pieces in prose and verse.* EETS O.S. 26. London: K. Paul, Trench, Trübner & Co. Third edition (first edition 1867, second edition 1914).	Religious treatise	This is a northern version of a translation of the Latin *Speculum S. Edmundi* by Edmund Rich, Archbishop of Canterbury, at Pontigny, around 1350, possibly from a Southern exemplar. *The Mirror of St. Edmund* (Vernon ms.) is another version of this text.	Translation from Latin

cmgaytry.m34	1350-1420	North	Dan Jon Gaytryge's Sermon	Perry, George G. 1969. *Dan Jon Gaytryge's sermon.* In George G. Perry (ed.), *Religious pieces in prose and verse.* EETS O.S. 26. New York: K. Paul, Trench, Trübner & Co. Third edition (first edition 1867, second edition 1914).	Sermon	The sermon is a translation made in 1357 by John Gaytryge, a monk at St. Mary's Abbey, York, of Archbishop Thoresby of York's Latin catechism. The English version is much longer than the Latin original. The *Lay Folks' Catechism* is a slightly different translation of the same work.	Translation from Latin
cmthorn.mx4	1420-1500#	North	Liber de Diversis Medicinis	Ogden, Margaret S. 1969. *The `Liber de diversis medicinis' in the Thornton Manuscript.* EETS O.S. 207. London: Oxford University Press. Originally published in 1938 (for 1936).	Handbooks, medicine		
cmpeterb.m1	1150-1250	East Midl.	The Peterborough Chronicle	Clark, Cecily. 1970. *The Peterborough Chronicle 1070-1154*. Oxford: Clarendon. Second edition (first edition 1958).	History		
cmvices1.m1	1150-1250	East Midl.	Vices and Virtues	Holthausen, Ferdinand. 1888. *Vices and virtues. Part 1.* EETS O.S. 89. London: Trübner.	Religious treatise		
cmearlps.m2	1250-1350	East Midl.	The Earliest Complete English Prose Psalter	Bülbring, Karl D. 1891. *The earliest complete English prose psalter.* EETS O.S. 97. London: K. Paul, Trench, Trübner & Co.	Bible		Translation from Latin /French
cmcloud.m3	1350-1420	East Midl.	The Cloud of Unknowing	Hodgson, Phyllis. 1944 (for 1943). *The cloud of unknowing and The book of privy counselling*. EETS O.S. 218. London: Oxford University Press	Religious treatise		

cmvices4.m34	1350-1420	East Midl.	The Book of Vices and Virtues	Francis, Winthrop N. 1942. *The book of vices and virtues: A fourteenth century English translation of the Somme le roi of Lorens D'Orléans*. EETS O.S. 217. London: Oxford University Press.	Religious treatise		Translation from French
cmreynes.m4	1420-1500	East Midl.	The Commonplace Book of Robert Reynes	Rother - Louis, Cameron. 1980. *The commonplace book of Robert Reynes of Acle: An edition of Tanner Ms. 407.* Garland Medieval Texts 1. New York: Garland.	Handbooks		
cmlamb1.m1	1150-1250	West Midl.	The Lambeth Homilies	Morris, Richard. 1969. *Old English homilies and homiletic treatises. Part I.* EETS O.S. 29, 34. New York: Greenwood Press. Originally published by Trübner (London, 1868).	Homilies		
cmancriw-1.m1	1150-1250	West Midl.	Ancrene Riwle	Ackerman, Robert W. and Roger Dahood. 1984. *Ancrene riwle. Introduction and Part I.* Medieval and Renaissance Texts and Studies 31. Binghamton, NY: Center for Medieval and Early Renaissance Studies, State University of New York at Binghamton.	Religious treatise		
cmancriw-2.m1	1150-1250	West Midl.		- || -			

cmaelr3.m23	1250-1350	West Midl.	Aelred of Rievaulx's De Institutione Inclusarum (Ms. Vernon)	Ayto, John and Alexandra Barratt. 1984. *Aelred of Rievaulx's De institutione inclusarum: Two English versions*. EETS O.S. 287. London: Oxford University Press.	Rule		Translation from Latin
cmedvern.m3	1350-1420	West Midl.	The Mirror of St. Edmund (Vernon Ms.)	Horstman, C. 1895-1896. *Yorkshire writers: Richard Rolle of Hampole*. London: Swan Sonnenschein and Co.	Religious treatise	A version of the Latin *Speculum S. Edmundi*. It is a closer rendering of the Latin than the *Thornton Ms.* version of the same text, also included in the corpus.	
cmmirk.m34	1350-1420	West Midl.	Mirk's Festial	Erbe, Theodore. 1905. *Mirk's Festial: A collection of homilies, by Johannes Mirkus (John Mirk). Part I*. EETS E.S. 96. London: K. Paul, Trench, Trübner & Co.	Sermons		
cmmalory.m4	1420-1500	West Midl.	Malory's Morte Darthur	Vinaver, Eugène. 1954. *The works of Thomas Malory*. London: Oxford University Press.	Romance		
cmkentho.m1	1150-1250	Kent	Kentish Homilies	Warner, Rubie D.-N. 1917 (for 1915). *Early English homilies from the twelfth-century ms. Vespasian D XIV*. EETS O.S. 152. London: K. Paul, Trench, Trübner & Co. Reprinted 1971 (publisher unknown).	Homilies		Translation from Latin
cmkentse.m1	1250-1350	Kent	Kentish Sermons	Hall, Joseph. 1963. *Selections from Early Middle English 1130-1250. Part I*. Oxford: Clarendon. Second edition (first edition 1920).	Homilies		Translation from French versions of Latin sermons

cmayenbi.m2	1250-1350	Kent	Ayenbite of Inwyt	Morris, Richard. 1979. *Dan Michel's Ayenbite of inwyt.* EETS O.S. 278. London: Oxford University Press. Originally published by Trübner (London, 1866) as EETS O.S. 23.	Religious treatise		Translation from French
cmpolych.m3	1350-1420	South	John of Trevisa's Polychronicon	Lumby, Joseph R. 1876, 1882. *Polychronicon Ranulphi Higden, monachi cestrensis, Vols. VI, VIII, English translations of John Trevisa and of an unknown writer of the fifteenth century.* Rolls Series 41. London: [publisher unknown].	History		Translation from Latin
cmntest.m3	1350-1420	South	The New Testament (Wycliffite)	Forshall, Josiah and Frederic Madden. 1879. *The New Testament in English according to the version of John Wycliffe about A.D. 1380 and revised by John Purvey about A.D. 1388.* Oxford: Clarendon	Bible	Dialect information taken from Helsinki Corpus -on the basis of LALME (PPCME2 team gives East Midlands)	Translation from Latin
cmpurvey.m3	1350-1420	South	Purvey's General Prologue to the Bible	Forshall, Josiah and Frederic Madden. 1850. *The Holy Bible, containing the Old and New Testaments, with the apocryphal books, in the earliest English versions made from the Latin Vulgate by John Wycliffe and his followers, Vol. 1.* Oxford: Oxford University Press. Reprinted 1982 (New York: AMS Press).	Religious treatise		

cmhorses.m3	1350-1420	South	A Late Middle English Treatise on Horses	Svinhufvud, Anne Charlotte. 1978. *A Late Middle English treatise on horses*. Stockholm Studies in English 47. Stockholm: Almqvist and Wiksell.	Handbooks, medicine		
cmroyal.m34	1350-1420*	South	Middle English Sermons	Ross, Woodburn O. 1940 (for 1938). *Middle English sermons edited, from British Museum ms. Royal 18 B. xxiii.* EETS O.S. 209. London: Oxford University Press.	Sermons		
cmgregor.m4	1420-1500	South	Gregory's Chronicle	Gairdner, James. 1876. *The historical collections of a citizen of London in the fifteenth century.* Camden Society, N.S. XVII. Westminster: Camden Society.	History		

APPENDIX 1C

*Text information for PPCME2 – Texts used for repeated sampling and genre groups (PTC) comparisons (***East Midlands** *sector)*

code	date	title	details (edition)	genre	additional info	translation
cmtrinit.mx1	1150-1250	Trinity Homilies	Morris, Richard. 1873. *Old English homilies of the twelfth century*. Second series. EETS O.S. 53. London: Trübner.	Homilies		Translation from Latin
cmctpars.m3	1350-1420	The parson's Tale	Benson, Larry D. 1987. The Riverside Chaucer. Boston: Houghton Mifflin. Third edition	Religious treatise		Translation from French
cmctmeli.m3	1350-1420	Tale of melibee	Benson, Larry D. 1987. *The Riverside Chaucer*. Boston: Houghton Mifflin. Third edition.	Philosophy/fiction		Translation from French
cmequato.m3	1350-1420	The Equatorie of the Planets	Price, Derek J. 1955. *The equatorie of the planetis*. Cambridge: Cambridge University Press.	Handbook/astronomy		Translation from Latin
cmwycser.m3	1350-1420	English Wycliffite Sermons	Hudson, Anne. 1983. *English Wycliffite sermons*. Oxford: Clarendon.	Sermons		
cmboeth.m3	1350-1420	Boethius	Benson, Larry D. 1987. *The Riverside Chaucer*. Boston: Houghton Mifflin. Third edition	Philosophy		Translation from Latin / French

cmotest.m3	1350-1420	The Old Testament (Wycliffe)	Forshall, Josiah and Frederic Madden. 1850. *The Holy Bible, containing the Old and New Testaments, with the apocryphal books, in the earliest English versions made from the Latin Vulgate by John Wycliffe and his follow-ers*, Vol. 1. Oxford: Oxford University Press. Reprinted 1982 (New York: AMS Press).	Bible		Translation from Latin
cmmandev.m3	1350-1420	Mandeville's Travels	Hamelius, Paul. 1919-1923 (for 1916). *Mandeville's travels, translated from the French of Jean D'Outremeuse*. EETS O.S. 153, 154. London: K. Paul, Trench, Trübner & Co.	Travelogue		Translation from French
cmastro.m3	1350-1420	Treatise on the Astrolabe	Benson, Larry D. 1987. *The Riverside Chaucer.* Boston: Houghton Mifflin. Third edition	Handbook/astronomy		
cmhilton.m34	comp. 1350-1420 ms date 1420-1500	Hilotn's Eight Chapters on Perfection	Kuriyagawa, Fumio. 1967. *Walter Hilton's Eight chapters on perfection*. Tokyo: Keio Institute of Cultural and Linguistic Studies.	Religious treatise		
cmjulnor.m34	comp. 1350-1420 ms date 1420-1500	Julian of Norwich's Revelations of Divine Love	Beer, Frances. 1978. *Julian of Norwich's revelations of divine love*: The shorter version edited from B.L. Add. Ms 37790. Middle English Texts 8. Heidelberg: Winter.	Religious treatise		

cmaelr4.m4	1420-1500	Aelred of Rievaulx's De Institutione Inclusarum	Ayto, John and Alexandra Barratt. 1984. *Aelred of Rievaulx's De institutione inclusarum*: Two English versions. EETS O.S. 287. London: Oxford University Press.	Religious treatise		Translation from French
cmedmund.m4	1420-1500	Life of St Edmund	Blake, Norman F. 1972. *The life of St. Edmund*. In Norman F. Blake (ed.), Middle English religious prose. York Medieval Texts. London: Arnold.	Biographies/Life of saints		
cmkempe.m4	1420-1500	The Book of Mary Kempe	Meech, Sanford B. and Hope E. Allen. 1940. The Book of Margery Kempe, Vol. 1. EETS O.S. 212. London: Oxford University Press.	Religious treatise		
cmcapser.m4	1420-1500	Capgrave's Sermon	Munro, John J. 1910. *John Capgrave's Lives of St. Augustine and St. Gilbert of Sempringham, and a sermon*. EETS O.S. 140. New York: K. Paul, Trench, Trübner & Co.	Sermons		
cmcapchr.m4	1420-1500	Capgrave's Chronicle	Lucas, Peter J. 1983. *John Capgrave's Abbreuiacion of cronicles*. EETS O.S. 285. Oxford: Oxford University Press	History		
cmreynar.m4	1420-1500	Caxton's History of Reynard the Fox	Blake, Norman F. 1970. *The History of Reynard the fox. Translated from the Dutch original by William Caxton*. EETS O.S. 263. London: Oxford University Press.	Fiction		Translation from Dutch

cmfitzja.m4	1420-1500	Fitzjames' Sermo die Lune	Jenkinson, Francis J.H. 1907. *Sermo die lune in ebdomada Pasche, by Richard Fitz-James*. Printed at Westminster by Wynkyn de Worde about the year 1495. Cambridge: Cambridge University Press. Facsimile edition.	Sermons		
cminnoce.m4	1420-1500	In Die Innocencium	Nichols, J.G. 1875. *Two sermons preached by the boy bishop, at St. Paul's Temp. Henry VII, and at Gloucester Temp. Mary*. Camden Society Miscellany VII. Camden Society N.S. XIV. London: [publisher unknown].	Sermons		

APPENDIX 2A

The content of the definition files for YCOE, listing elements included as an object, finite verb and nonfinite verb.
non_finite_verb: *VB|*VBN*|*VAG*|*HV|*HVN*|*HAG*|*BE|*BEN*
finite_verb: *MDP*|*MDD*|*HVP*|*HVD*|*BEP*|*BED*|*VBP*|*VBD*|*AXD*|*AXP*
object: NP|NP-ACC|NP-GEN|NP-DAT|NP-RFL|NP-ACC-RFL|NP-GEN-RFL|NP-DAT-RFL|NP-RSP|NP-ACC-RSP|NP-GEN-RSP|NP-DAT-RSP
Tags explained:
-NOM, -ACC, -DAT, -GEN – tag extension indicating cases (otherwise other grammatical function) at the phrase level
AXD – auxiliary verb, infinitive
AXP – auxiliary verb, present indicative/subjunctive ambiguous
BE – BE, infinitive (inflected) (^D)
BEP – BE, present indicative/subjunctive ambiguous
BED – BE, past indicative/subjunctive ambiguous
BEN – BE, past participle (^N, ^A, ^G, ^D)
HAG – HAVE, present participle (^N, ^D, ^G)
HV – HAVE, infinitive (inflected)(^D)
HVD – HAVE, past indicative/subjunctive
HVN – HAVE, past participle (verbal or adjectival) (^N)
HVP – HAVE, present indicative/subjunctive
IP-MAT – matrix clause
IP-SUB – subordinate clause
MDD – modal verb, past indicative/subjunctive ambiguous
MDP – modal verb, present indicative/subjunctive ambiguous
NP - noun phrase, argument
NP-ACC - noun phrase in Accusative
NP-DAT - noun phrase in Dative
NP-GEN - noun phrase in Genitive
NP-NOM – noun phrase subject in tensed clauses
NP-RFL – noun phrase, reflexive argument
NP-ACC-RFL - noun phrase, reflexive argument in Accusative
NP-DAT-RFL - noun phrase, reflexive argument in Dative
NP-GEN-RFL - noun phrase, reflexive argument in Genitive
NP-RSP – noun phrase argument, resumptive
NP-ACC-RSP – noun phrase argument, resumptive in Accusative
NP-DAT-RSP – noun phrase argument, resumptive in Dative
NP-GEN-RSP – noun phrase argument, resumptive in Genitive
VAG – lexical verb, present participle)
VB – lexical verb, infinitive (inflected) (^D)
VBD – lexical verb, past tense indicative/subjunctive ambiguous
VBN – past participle (^N, ^A, ^G, ^D, ^I)
VBP – lexical verb, present indicative/subjunctive ambiguous

APPENDIX 2B

The content of the definition files for PPCME2, listing elements included as finite verb and nonfinite verb, and non-pronominal NP. Objects are not tagged for cases in PPCME2, no separate definition file is needed.

finite_verb: *MD|*HVP|*HVD|*DOP|*DOD|*BEP|*BED|*VBP|*VBD
non_finite_verb: *VB|V*N|*HV|H*N|*DO|D*N|*BE|BEN
non-pronominal_NP: *N*|D*|Q*|ADJ*|CONJ*|*ONE*|*OTHER*

Tags explained:
ADJ – adjective
BE – BE, infinitive
BED – BE, past (including past subjunctive)
BEN – BE, perfect participle
BEP – BE, present (including present subjunctive)
CONJ – coordinating conjunction
D – determiner
D*N – DO, passive/perfect participle (verbal or adjectival)
DO – DO, infinitive
DOD – DO, past (including past subjunctive)
DOP – DO, present (including present subjunctive)
H*N – HAVE, passive/perfect participle (verbal or adjectival)
HV – HAVE, infinitive
HVD – HAVE, past (including past subjunctive)
HVP – HAVE, present (including present subjunctive)
IP-MAT – matrix clause
IP-SUB – subordinate clause
MD – modal verb
N – noun
NP-SBJ – noun phrase subject
NP-OB1, NP-OB2 – noun phrase (first, second) object
ONE – the word 'one'
OTHER – the word 'other'
Q – Quantifier
VBP – verb (other than BE, DO, HV) present (including present subjunctive)
VBD – verb (other than BE, DO, HV) past (including past subjunctive)
VB – infinitive, verbs other than (BE, DO HV)
V*N – passive/perfect participles (for verb other than BE, DO, HV) (verbal or adjectival)

APPENDIX 3A

Examples of retrieved sequences, with query line marked in bold type. Each example begins with plain text version followed by a parsed one.

YCOE:

SVO with a single finite verb:

query: ((IP-MAT*/SUB* iDoms NP-NOM*)
AND (NP-NOM* iPrecedes finite_verb)
AND (finite_verb iPrecedes object))

```
/~*
Ond he dyde monig heofonlic wundor,
*~/

( (IP-MAT (CONJ Ond)
                          (NP-NOM (PRO^N he))
                                (VBD dyde)
                          (NP-ACC (Q^A monig)
                                                (ADJ^A heofonlic)
                                                (N^A wundor)
                                                      (, ,)
(ID comart3,Mart_5_[Kotzor]:Jy27,A.25.1289))

/~*
Cristes leorningcnihtas leidon heora reaf uppon +tam assan. for +dan +te he
nolde. on nacedum assan ridan.
(cocathom1,+ACHom_I,_14.1:293.98.2641)
*~/

 ( (IP-MAT (NP-NOM (NP-GEN (NR^G Cristes))
                                                (N^N leorningcnihtas))
                                      (VBDI leidon)
                  (NP-ACC (PRO$ heora) (N^A reaf))
                                      (PP (P uppon)
                  (NP-DAT (D^D +tam) (N^D assan)))
                                                (, .)
                                      (CP-ADV (P for)
                                                (D^I +dan)
                                                  (C +te)
                                                (IP-SUB RMV:he_nolde_....))
                                              (. .))
  (ID cocathom1,+ACHom_I,_14.1:293.98.2641))

/~*
and +ta Deniscan ahton w+alstowe gewealde.
(cochronC,ChronC_[Rositzke]:833.1.507)
*~/

 ( (IP-MAT (CONJ and)
         (NP-NOM (D^N +ta) (ADJ^N Deniscan))
                                      (VBDI ahton)
              (NP-ACC (NP-GEN (N^G w+alstowe))
                                                (N^A geweald))
                                              (. .))
  (ID cochronC,ChronC_[Rositzke]:833.1.507))
```

SVO with finite + nonfinite:

query: ((IP-MAT*/SUB* iDoms NP-NOM*)
AND (NP-NOM* iPrecedes finite_verb)
AND (finite_verb iPrecedes non_finite_verb)
AND (non_finite_verb iPrecedes object))

```
/~*
We willa+d secgan eow sum byspel.
(cocathom1,+ACHom_I,_14.1:294.111.2653)
*~/

 ( (IP-MAT (NP-NOM (PRO^N We))
                                    (MDPI willa+d)
                                      (VB secgan)
                                    (NP (PRO eow))
                (NP-ACC (Q^A sum) (N^A byspel))
                                                (. .))
  (ID cocathom1,+ACHom_I,_14.1:294.111.2653))


/~*
& +d+at heafod sceal wisian +d+am fotum, +d+at hie st+appen on ryhtne weg:
(cocura,CP:18.131.22.896)
*~/

 ( (IP-MAT (CONJ &)
           (NP-NOM (D^N +d+at) (N^N heafod))
                                    (MDPI sceal)
                                     (VB wisian)
            (NP-DAT (D^D +d+am) (N^D fotum))
                                             (, ,)
                                (CP-THT (C +d+at)
                                                    (IP-SUB RMV:hie_st+appen_on...))
                                             (. :))
  (ID cocura,CP:18.131.22.896))


/~*
& hire sweostor ges+at big H+alendes fotum, +t+are nama w+as Maria; for+ton
+te heo wolde gehyran his word & his lare.
(coblick,HomS_21_[BlHom_6]:67.30.825)
*~/

( (IP-SUB (NP-NOM (PRO^N heo))
                                    (MDD wolde)
                                    (VB gehyran)
       (NP-ACC (NP-ACC (PRO$ his) (N^A word))
                                                 (CONJP (CONJ &)

  (NP-ACC (PRO$ his) (N^A lare)))))
   (ID coblick,HomS_21_[BlHom_6]:67.30.825))
```

PPCME2:

SVO with a single finite verb:

query: ((IP-MAT*/SUB* iDoms NP-SBJ*)
AND (NP-SBJ* iPrecedes finite_verb)
AND (finite_verb iPrecedes NP-OB*))

```
/~*
god +treate+d +teose +turch ysaye.
(CMANCRIW-1,II.160.2181)
*~/

( (IP-MAT (NP-SBJ (NPR god))
                                    (VBP +treate+d)
                                (NP-OB1 (D +teose))
                                    (PP (P +turch)
                                (NP (NPR ysaye)))
                                        (. .))
 (ID CMANCRIW-1,II.160.2181))


/~*
Our Lord herd +te desire of +te $pouer in gost;
(CMEARLPS,11.403)
*~/

( (IP-MAT (NP-SBJ (PRO$ Our) (NPR Lord))
                                        (VBD herd)
                                (NP-OB1 (D +te)
                                                (N desire)
                                                (PP (P of)
                                                  (NP (D +te)
                                                        (ADJ $pouer)
                                                        (PP (P in)
                                                          (NP (N gost))))))
                                        (. ;))
 (ID CMEARLPS,11.403))


/~*
+Tis hope conforted me in my meknes,
(CMEARLPS,148.6536)
*~/

( (IP-MAT (NP-SBJ (D +Tis) (N hope))
                                    (VBD conforted)
                                (NP-OB1 (PRO me))
                                        (PP (P in)
                            (NP (PRO$ my) (N meknes)))
                                        (. ,))
 (ID CMEARLPS,148.6536))


/~*
And this man was Adam, by whom synne entred into this world, whan he brak the
comaundementz of God.
(CMCTPARS,296.C2b.356)
*~/
```

```
( (IP-SUB (NP-SBJ (PRO he))
                                        (VBD brak)
                                  (NP-OB1 (D the)
                                                (NS comaundementz)
                                                (PP (P of)
                                                  (NP (NPR God)))))
 (ID CMCTPARS,296.C2b.356))


/~*
veray God yn +tat he turned watyr ynto wyne, and veray man yn +tat he ete and
drangke wyth hom.
(CMMIRK,52.1469)
*~/

( (IP-SUB (NP-SBJ (PRO he))
                                      (VBD turned)
                                (NP-OB1 (N watyr))
                                      (PP (P ynto)
                                    (NP (N wyne))))
 (ID CMMIRK,52.1469))
```

SVO with finite + nonfinite verb:

query: ((IP-MAT*/SUB* iDoms NP-SBJ*)
AND (NP-SBJ* iPrecedes finite_verb)
AND (finite_verb iPrecedes non_finite_verb)
AND (non_finite_verb iPrecedes NP-OB*))

```
/~*
ye sal do o+ter mens wille, yef ye wil cum to god.
(CMBENRUL,10.330)
*~/

 ( (IP-MAT (NP-SBJ (PRO ye))
                                        (MD sal)
                                        (DO do)
 (NP-OB1 (NP-POS (OTHER o+ter) (NS$ mens))
                                                (N wille))
                                          (, ,)
                                      (PP (P yef)
                                    (CP-ADV (C 0)
                                                  (IP-SUB RMV:ye_wil_cum...)))
                                          (. .))
 (ID CMBENRUL,10.330))


/~*
and +te Britons wolde nou+gt +geue here dou+gtres to +to Straungers; Wherfore
+tai went ouere +te see into Irland, and brou+gt with ham Wymmen, and ham +to
spousede;
(CMBRUT3,36.1124)
*~/

 ( (IP-MAT (CONJ and)
                (NP-SBJ (D +te) (NPRS Britons))
```

(MD wolde)
(NEG nou+gt)
(VB +geue)
(NP-OB1 (PRO$ here) (NS dou+gtres))
(PP (P to)
(NP (D +to) (NS Straungers)))
(, ;)
(CP-CAR (WPP-1 (WADV+P Wherfore))
(C 0)
(IP-SUB RMV:*T*-1_+tai_went...))
(. ;))
(ID CMBRUT3,36.1124))

/~*
and horde hir there, that no man shold perceyve hir grete sorowys.
(CMMALORY,637.3858)
*~/

((IP-SUB (NP-SBJ (Q no) (N man))
(MD shold)
(VB perceyve)
(NP-OB1 (PRO$ hir) (ADJ grete) (NS sorowys)))
(ID CMMALORY,637.3858))

APPENDIX 3B

Examples of retrieved sequences that include empty categories (PPCME2)

Subject elided under conjunction [*con*]:

```
/~*
and helde +te oynement on Ihesu heed.
(CMAELR3,44.541)
*~/

 ( (IP-MAT (CONJ and)
                                    (NP-SBJ *con*)
                                    (VBD helde)
                          (NP-OB1 (D +te) (N oynement))
                                    (PP (P on)
                             (NP (NPR$ Ihesu) (N heed)))
                                    (. .))
  (ID CMAELR3,44.541))
```

Empty expletive subject (in impersonal constructions, complement clauses, among others) [*exp*]:

```
/~*
and by+tou+gt hym +tat tweyne of +te kynges schippes were to-broke wi+t
tempest in +te see, and i-spoyled by men of Cipres.
(CMPOLYCH,VIII,107.3709)
*~/

 ( (IP-MAT (CONJ and)
                                 (NP-SBJ-1 *exp*)
                                 (VBD by+tou+gt)
                               (NP-OB2 (PRO hym))
                                 (CP-THT-1 (C +tat)
                                                  (IP-SUB RMV:tweyne_of_+te...))
                                 (. .))
  (ID CMPOLYCH,VIII,107.3709))
```

Other empty subjects (including pro-drop) [*pro*]:

```
/~*
and bare a sone +tat was callede Arthure.
(CMBRUT3,67.2022)
*~/

( (IP-MAT (CONJ and)
                                 (NP-SBJ *pro*)
                                   (VBD bare)
                                 (NP-OB1 (D a)
                                                 (N sone)

                               (CP-REL (WNP-1 0)
                                                   (C +tat)
                                                   (IP-SUB RMV:*T*-1_was_callede...)))
                                 (. .))
  (ID CMBRUT3,67.2022))
```

Search functions, logical operators as well as search control commands and output format commands apply to both corpora in the same manner.

APPENDIX 4

Table illustrating the growth rate of the preference for SVO. The percentages represent differences between the (higher) average normed frequencies recorded in the ME dialect sectors and their corresponding, (lower) OE ones. (The values found in Figure 6.1.1a)

Freq. norm.	**A NM - N**	**A M - EM**	**A M - WM**	**K - K**	**WS - S**
Including empty categories					
25IP	72%	68%	64%	56%	59%
100 IP	72%	63%	59%	50%	60%
200 IP	79%	63%	59%	50%	61%
400 IP	82%	59%	55%	72%	60%
Excluding empty categories					
25 IP	71%	68%	64%	55%	61%
100 IP	71%	63%	58%	49%	62 %
200 IP	82%	63%	58%	52%	64%
400 IP	84%	58%	52%	73%	63%
Overall					

APPENDIX 5A

Query delimiting and refining: Estimate for PPCME2 on discontinuous VP:

IP-MAT

	BEFORE CORRECTION				Tokens with discont/ unit		AFTER CORRECTION (SUBTRACTION)			
	f/25 wt	**f/25 nt**	**CV wt**	**CV nt**	**Tkns wt**	**Tkns nt**	**f/25 wt**	**f/25 nt**	**CV wt**	**CV nt**
N	8.02	8.61	25	21	9	8	7.98	8.56	26	21
EM	6.87	7.98	38	26	32	32	6.68	7.78	38	27
WM	6.88	8.27	14	8	65	59	6.73	8.10	14	9
K	7.10	7.55	10	12	26	26	6.94	7.42	9	10
S	7.59	9.23	17	14	8	6	7.55	9.18	16	14

IP-SUB

	BEFORE CORRECTION				Tokens with discont/ unit		AFTER CORRECTION (SUBTRACTION)			
	f/25 wt	**f/25 nt**	**CV wt**	**CV nt**	**Tkns wt**	**Tkns nt**	**f/25 wt**	**f/25 nt**	**CV wt**	**CV nt**
N	4.33	6.53	27	17	23	16	4.27	6.47	26	16
EM	3.88	6.36	23	30	56	44	3.72	6.16	25	33
WM	3.42	5.45	48	39	81	55	3.25	5.20	53	43
K	3.30	5.32	22	18	64	49	3.01	4.93	31	21
S	3.79	6.05	8	21	18	9	3.72	5.92	10	24

wt – sequences with empty categories
nt – sequences without empty categories

APPENDIX 5B

Query delimiting and refining: Comparison of results for PPCME2 on searches which include and exclude ADVP; sequences without the empty categories:

SVO in IP-MAT	EXCLUDING ADVP			INCLUDING ADVP			DIFFERENCE		
	Tokens	F/25	CV	Tokens	F/25	CV	Tokens (%)	F/25	CV
N	988	7.98	26	1012	8.16	24	24 (2%)	0.18	2
EM	2120	6.68	38	2147	6.82	37	27 (1%)	0.14	1
WM	3481	6.73	14	3548	6.87	14	67 (2%)	0.14	-
K	808	6.94	9	827	7.27	5	19 (2%)	0.33	4
S	2307	7.55	16	2377	7.78	17	70 (3%)	0.23	1

SVO in IP-SUB	EXCLUDING ADVP			INCLUDING ADVP			DIFFERENCE		
	Tokens	F/25	CV	Tokens	F/25	CV	Tokens (%)	F/25	CV
N	999	4.27	26	1032	4.42	25	33 (3%)	0.15	1
EM	721	3.72	25	741	3.79	24	20 (3%)	0.07	1
WM	2179	3.25	53	2232	3.33	52	53 (2%)	0.08	1
K	453	3.01	31	479	3.14	29	26 (5%)	0.13	2
S	828	3.72	10	851	3.80	11	23 (3%)	0.08	1

SVO in IP-	EXCLUDING ADVP			INCLUDING ADVP			DIFFERENCE		
	Tokens	F/25	CV	Tokens	F/25	CV	Tokens (%)	F/25	CV
N	1987	5.56	6	2044	5.72	6	57 (3%)	0.16	-
EM	2841	5.24	35	2888	5.35	34	47 (2%)	0.11	1
WM	5660	4.61	34	5780	4.71	33	120 (2%)	0.1	1
K	1261	4.62	15	1306	4.83	13	45 (3%)	0.21	2
S	3135	5.60	8	3228	5.75	9	93 (3%)	0.15	1

APPENDIX 6

Distribution of SVO for PPCME2 when excluding translations for the common-base of 200 clauses (200 IP):

F/200 IP	SVO IP-	SVO IP- no transl.	SVO IP-MAT	SVO IP-MAT no transl.	SVO IP-SUB	SVO IP-SUB no transl.
Sequencing including empty categories						
N	58.20	57.30	60.05	58.25	54.72	56.88
EM	55.44	50.21	62.28	56.05	49.29	43.33
WM	50.33	50.43	64.80	65.79	41.63	42.28
K	47.62	n/a	57.10	n/a	39.44	n/a
S	59.87	57.26	73.54	72.60	47.40	45.00
Sequencing excluding empty categories						
N	44.88	45.26	54.87	54.17	35.99	39.97
EM	41.90	36.99	53.42	44.42	29.73	31.11
WM	36.87	35.77	53.81	50.26	26.00	27.51
K	36.95	n/a	54.62	n/a	24.05	n/a
S	44.78	44.65	57.19	57.12	29.77	29.02

APPENDIX 7A

Repeated sampling procedures – sample rounds with the CV calculated before and after the extraction of translations (Expanded version of table 6.1.2.3d)

Sample symbol	Texts used: Bold type - not translations
A	**cmpeterb.m1** cmtrinit.mx1 cmearlps.m2 cmctpars.m3 **cmwycser.m3** **cmcapser.m4**
B	**cmpeterb.m1** cmvices1.m1 cmearlps.m2 cmctpars.m3 **cmastro.m3** **cmkempe.m4**
C	**cmpeterb.m1** cmtrinit.mx1 cmearlps.m2 cmctpars.m3 **cmjulnor.m34** **cmcapser.m4**
D	**cmpeterb.m1** cmvices1.m1 cmearlps.m2 cmctmeli.m3 **cmwycser.m3** **cmcapchr.m4**
E	**cmpeterb.m1** cmtrinit.mx1 cmearlps.m2 cmctmeli.m3 **cmastro.m3** **cmreynes.m4**
F	**cmpeterb.m1** cmvices1.m1 cmearlps.m2 cmctmeli.m3 **cmjulnor.m34** **cmfitzja.m4**
G	**cmpeterb.m1** cmtrinit.mx1 cmearlps.m2 cmequato.m3 **cmwycser.m3** **cminnoce.m4**
H	**cmpeterb.m1** cmvices1.m1 cmearlps.m2 cmequato.m3 **cmastro.m3** **cmcapchr.m4**

I	**cmpeterb.m1** cmtrinit.mx1 cmearlps.m2 cmequato.m3 **cmjulnor.m34** **cmcapchr.m4**
J	**cmpeterb.m1** cmvices1.m1 cmearlps.m2 cmboeth.m3 **cmastro.m3** **cminnoce.m4**
K	**cmpeterb.m1** cmtrinit.mx1 cmearlps.m2 cmboeth.m3 **cmjulnor.m34** **cmkempe.m4**
L	**cmpeterb.m1** cmvices1.m1 cmearlps.m2 cmboeth.m3 **cmwycser.m3** **cmkempe.m4**
M	**cmpeterb.m1** cmtrinit.mx1 cmearlps.m2 cmotest.m3 **cmastro.m3** **cmfitzja.m4**
N	**cmpeterb.m1** cmvices1.m1 cmearlps.m2 cmotest.m3 **cmcloud.m3** **cmcapser.m4**
O	**cmpeterb.m1** cmtrinit.mx1 cmearlps.m2 cmotest.m3 **cmcloud.m3** **cmkempe.m4**
P	**cmpeterb.m1** cmvices1.m1 cmearlps.m2 cmmandev.m3 **cmcloud.m3** **cminnoce.m4**
Q	**cmpeterb.m1** cmtrinit.mx1 cmmandev.m3 cmotest.m3 **cmcloud.m3** **cmcapchr.m4**

R	**cmpeterb.m1** cmvices1.m1 cmearlps.m2 cmmandev.m3 **cmhilton.m34** **cmcapchr.m4**
S	**cmpeterb.m1** cmtrinit.mx1 cmearlps.m2 cmvices4.m34 **cmhilton.m34** **cminnoce.m4**
T	**cmpeterb.m1** cmvices1.m1 cmearlps.m2 cmvices4.m34 **cmhilton.m34** **cmkempe.m4**
U	**cmpeterb.m1** cmtrinit.mx1 cmearlps.m2 cmvices4.m34 **cmhilton.m34** **cmfitzja.m4**

APPENDIX 7B

Repeated sampling procedures – samples with a uniform distribution of texts in terms of subperiodisation (Expanded version of tables 6.1.2.4b and c)

Each sample follows the sample subperiod layout:
1150-1250
1150-1250
1250-1350
1350-1420
1350-1420
1420-1500

Sample no.	Texts used
1	cmtrinit.mx1 cmvices1.m1 cmearlps.m2 cmcloud.m3 cmctpars.m3 cmkempe.m4
2	cmtrinit.mx1 cmvices1.m1 cmearlps.m2 cmwycser.m3 cmhilton.m34 cmaelr4.m4
3	cmtrinit.mx1 cmvices1.m1 cmearlps.m2 cmmandev.m3 cmjulnor.m34 cmfitzja.m4
4	cmtrinit.mx1 cmvices1.m1 cmearlps.m2 cmequato.m3 cmastro.m3 cmreynar.m4
5	cmtrinit.mx1 cmvices1.m1 cmearlps.m2 cmctmeli.m3 cmboeth.m3 cmkempe.m4
6	cmtrinit.mx1 cmvices1.m1 cmearlps.m2 cmmandev.m3 cmotest.m3 cmcapser.m4
7	cmtrinit.mx1 cmvices1.m1 cmearlps.m2 cmequato.m3 cmwycser.m3 cminnoce.m4

8	cmtrinit.mx1 cmpeterb.m1 cmearlps.m2 cmvices4.m34 cmjulnor.m34 cmaelr4.m4
9	cmtrinit.mx1 cmpeterb.m1 cmearlps.m2 cmmandev.m3 cmwycser.m3 cmkempe.m4
10	cmtrinit.mx1 cmpeterb.m1 cmearlps.m2 cmctmeli.m3 cmboeth.m3 cmfitzja.m4
11	cmtrinit.mx1 cmpeterb.m1 cmearlps.m2 cmequato.m3 cmastro.m3 cmreynes.m4
12	cmtrinit.mx1 cmpeterb.m1 cmearlps.m2 cmvices4.m34 cmotest.m3 cmkempe.m4
13	cmtrinit.mx1 cmpeterb.m1 cmearlps.m2 cmcloud.m3 cmctpars.m3 cminnoce.m4
14	cmtrinit.mx1 cmpeterb.m1 cmearlps.m2 cmctmeli.m3 cmcapser.m4 cmedmund.m4
15	cmtrinit.mx1 cmpeterb.m1 cmearlps.m2 cmboeth.m3 cmhilton.m34 cmaelr4.m4
16	cmtrinit.mx1 cmpeterb.m1 cmearlps.m2 cmcloud.m3 cmctpars.m3 cmkempe.m4

APPENDIX 8

Additional information for Table 6.1.2.4d: Comparison of standard samples with the sampling units containing texts from a single PTC, sequencing with and without empty categories: common base of 25 clauses

PPCME2 texts used in the standard sample comparisons:

Dialect	Filename
North	cmrolltr.m24 cmrollep.m24 cmbenrul.m3 cmedthor.m34 cmgaytry.m34 cmthorn.mx4
East Midlands	cmpeterb.m1 cmvices1.m1 cmearlps.m2 cmcloud.m3 cmvices4.m34 cmreynes.m34
West Midlands	cmlamb1.m1 cmancriw-1.m1 cmancriw-2.m1 cmaelr3.m23 cmedevern.m3 cmmirk.m34 cmmalory.m34
Kent	cmkentho.m1 cmkentse.m1 cmayenbi.m2
South	cmpolych.m3 cmpurvet.m3 cmhorses.m3 cmroyal.m34 cmgregor.m4

Texts used in the IR sample - after the extraction of texts from the standard sample that did not belong to IR PTC:

Dialect	Filename
North	cmrolltr.m24 cmrollep.m24 cmedthor.m34 cmgaytry.m34
East Midlands	cmvices1.m1 cmcloud.m3 cmvices4.m34
West Midlands	cmlamb1.m1 cmancriw-1.m1 cmancriw-2.m1 cmedvern.m3 cmmirk.m34
Kent	cmkentho.m1 cmkentse.m1 cmayenbi.m2
South	cmpurvey.m3 cmroyal.m34

PUBLICATIONS OF THE UNIVERSITY OF EASTERN FINLAND DISSERTATIONS IN EDUCATION, HUMANITIES, AND THEOLOGY

1. Taru Viinikainen. *Taipuuko "akrobaatti Aleksandra"? Nimikekonstruktio ja nimikkeen taipuminen lehtikielessä 1900-luvulta 2000-luvulle.* 2010.
2. Pekka Metso. *Divine Presence in the Eucharistic Theology of Nicholas Cabasilas.* 2010.
3. Pekka Kilpeläinen. *In Search of a Postcategorical Utopia. James Baldwin and the Politics of 'Race' and Sexuality.* 2010.
4. Leena Vartiainen. *Yhteisöllinen käsityö. Verkostoja, taitoja ja yhteisiä elämyksiä.* 2010.
5. Alexandra Simon-López. *Hypersurrealism. Surrealist Literary Hypertexts.* 2010.
6. Merja Sagulin. *Jälkiä ajan hiekassa. Kontekstuaalinen tutkimus Daniel Defoen Robinson Crusoen suomenkielisten adaptaatioiden aatteellisista ja kirjallisista traditioista sekä subjektikäsityksistä.* 2010.
7. Pirkko Pollari. *Vapaan sivistystyön kieltenopettajien pedagogiset ratkaisut ja käytänteet teknologiaa hyödyntävässä vieraiden kielten opetuksessa.* 2010.
8. Ulla Piela. *Kansanparannuksen kerrotut merkitykset Pohjois-Karjalassa 1800- ja 1900-luvuilla.* 2010.
9. Lea Meriläinen. *Language Transfer in the Written English of Finnish Students.* 2010.
10. Kati Aho-Mustonen. *Group Psychoeducation for Forensic Long-term Patients with Schizophrenia.* 2011.
11. Anne-Maria Nupponen. *»Savon murre» savolaiskorvin. Kansa murteen havainnoijana.* 2011.
12. Teemu Valtonen. *An Insight into Collaborative Learning with ICT: Teachers' and Students' Perspectives.* 2011.
13. Teemu Kakkuri. *Evankelinen liike kirkossa ja yhteiskunnassa 1944-1963. Aktiivinen uudistusliike ja konservatiivinen sopeutuja.* 2011.
14. Riitta Kärkkäinen. *Doing Better? Children's and Their Parents' and Teachers' Perceptions of the Malleability of the Child's Academic Competences.* 2011.
15. Jouko Kiiski. *Suomalainen avioero 2000-luvun alussa. Miksi avioliitto puretaan, miten ero koetaan ja miten siitä selviydytään.* 2011.
16. Liisa Timonen. *Kansainvälisty tai väisty? Tapaustutkimus kansainvälisyysosaamisen ja kulttuurienvälisen oppimisen merkityksenannoista oppijan, opettajan ja korkeakoulutoimijan pedagogisen suhteen rajaamissa kohtaamisen tiloissa.* 2011.
17. Matti Vänttinen. *Oikeasti hyvä numero. Oppilaiden arvioinnin totuudet ja totuustuotanto rinnakkaiskoulusta yhtenäiskouluun.* 2011.
18. Merja Ylönen. *Aikuiset opin poluilla. Oppimistukikeskuksen asiakkaiden opiskelukokemuksista ja kouluttautumishalukkuudelle merkityksellisistä tekijöistä.* 2011.
19. Kirsi Pankarinkangas. *Leskien keski-iässä tai myöhemmällä iällä solmimat uudet avioliitot. Seurantatutkimus.* 2011.
20. Olavi Leino. *Oppisopimusopiskelijan oppimisen henkilökohtaistaminen ja oppimismahdollisuudet työpaikalla.* 2011.
21. Kristiina Abdallah. *Translators in Production Networks. Reflections on Agency, Quality and Ethics.* 2012.

22. Riina Kokkonen. *Mittarissa lapsen keho ja vanhemmuus – tervettä lasta sekä ”hyvää” ja ”huonoa” vanhemmuutta koskevia tulkintoja nyky-Suomessa.* 2012.
23. Ari Sivenius. *Aikuislukion eetos opettajien merkityksenantojen valossa.* 2012.
24. Kamal Sbiri. *Voices from the Margin. Rethinking History, Identity, and Belonging in the Contemporary North African Anglophone Novel.* 2012.
25. Ville Sassi. *Uudenlaisen pahan unohdettu historia. Arvohistoriallinen tutkimus 1980-luvun suomalaisen romaanin pahan tematiikasta ja ”pahan koulukunta” –vuosikymmenmääritteen muodostumisesta kirjallisuusjärjestelmässä.* 2012.
26. Merja Hyytiäinen. *Integroiden, segregoiden ja osallistaen. Kolmen vaikeasti kehitysvammaisen oppilaan opiskelu yleisopetuksessa ja koulupolku esiopetuksesta toiselle asteelle.* 2012.
27. Hanna Mikkola. *”Tänään työ on kauneus on ruumis on laihuus.” Feministinen luenta syömishäiriöiden ja naissukupuolen kytköksistä suomalaisissa syömishäiriöromaaneissa.* 2012.
28. Aino Äikäs. *Toiselta asteelta eteenpäin. Narratiivinen tutkimus vaikeavammaisen nuoren aikuisen koulutuksesta ja työllistymisestä.* 2012.
29. Maija Korhonen. *Yrittäjyyttä ja yrittäjämäisyyttä kaikille? Uusliberalistinen hallinta, koulutettavuus ja sosiaaliset erot peruskoulun yrittäjyyskasvatuksessa.* 2012.
30. Päivikki Ronkainen. *Yhteinen tehtävä. Muutoksen avaama kehittämispyrkimys opettajayhteisössä.* 2012.
31. Kalevi Paldanius. *Eläinlääkärin ammatti-identiteetti, asiakasvuorovaikutuksen jännitteiden hallinta ja kliinisen päättelyn yhteenkietoutuminen sekapraktiikassa.* 2012.
32. Kari Korolainen. *Koristelun kuvailu. Kategorisoinnin analyysi.* 2012.
33. Maija Metsämäki. *Influencing through Language. Studies in L2 Debate.* 2012.
34. Pål Lauritzen. *Conceptual and Procedural Knowledge of Mathematical Functions.* 2012.
35. Eeva Raunistola-Juutinen. *Äiti ja nunna - Kirkkojen maailmanneuvoston naisten vuosikymmenen ortodoksiset naiskuvat.* 2012.
36. Marja-Liisa Kakkonen. *Learning Entrepreneurial Competences in an International Undergraduate Degree Programme. A Follow-Up Study.* 2012.
37. Outi Sipilä. *Esiliina aikansa kehyksissä – moniaikaista tekstiilikulttuuria ja representaatioita kodista, perheestä, puhtaudesta ja käsityöstä 1900-luvun alkupuolen Suomessa.* 2012.
38. Seija Jeskanen. *Piina vai pelastus? Portfolio aineenopettajaopiskelijoiden ammatillisen kehittymisen välineenä.* 2012.
39. Reijo Virolainen. *Evankeliumin asialla - Kurt Frörin käsitys evankelisesta kasvatuksesta ja opetuksesta Saksassa 1930-luvulta 1970-luvulle.* 2013.
40. Katarzyna Szal. *Finnish Literature in Poland, Polish Literature in Finland – Comparative Reception Study from a Hermeneutic Perspective.* 2013.
41. Eeva-Liisa Ahtiainen. *Kansainvälistymisen ja laadunvarmistuksen yhteys ammattikorkeakoulun asiakirjateksteissä. Tapaustutkimus.* 2013.
42. Jorma Pitkänen. *Fides Directa – Fides Reflexa. Jonas Laguksen käsitys vanhurskauttavasta uskosta.* 2013.
43. Riitta Rajasuu. *Kuopiossa, Oulussa ja Turussa vuosina 1725–1744 ja 1825–1844 syntyneiden kastenimet.* 2013.
44. Irina Karvonen. *Pyhän Aleksanteri Syväriläisen koulukunta – 1500-luvun luostarihistoriaa vai 1800-luvun venäläiskansallista tulkintaa?* 2013.

45. Meri Kytö. *Kotiin kuuluvaa. Yksityisen ja yhteisen kaupunkiäänitilan risteymät.* 2013.
46. Jörg Weber. *Die Idee von der Mystagogie Jesu im geistigen Menschen: Einführung in die »christliche Theosophie« des Corpus Areopagiticum.* 2013.
47. Tuija, Lukin. *Motivaatio matematiikan opiskelussa – seurantatutkimus motivaatiotekijöistä ja niiden välisistä yhteyksistä yläkoulun aikana.* 2013.
48. Virpi Kaukio. *Sateenkaari lätäkössä. Kuvitellun ja kerrotun ympäristöestetiikka.* 2013.
49. Susanna Pöntinen. *Tieto- ja viestintäteknologian opetuskäytön kulttuurin diskursiivinen muotoutuminen luokanopettajaopiskelijoiden puheessa.* 2013.
50. Maria Takala-Roszczenko. *The 'Latin' within the 'Greek': The Feast of the Holy Eucharist in the Context of Ruthenian Eastern Rite Liturgical Evolution in the 16th–18th Centuries.* 2013
51. Erkki Nieminen. *Henki vastaan alkoholi: AA-toiminnan synty ja kehitys Lahdessa 1950-1995.* 2014.
52. Jani Kaasinen. *Perinnerakentaminen käsitteenä ja osana teknologiakasvatusta - opettajaopiskelijoiden käsitykset, käsitysten jäsentyneisyys ja muutos perinnerakentamisen opintojakson aikana.* 2014.
53. Gerson Lameck Mgaya. *Spiritual gifts: A sociorhetorical interpretation of 1 cor 12–14.* 2014.
54. Pauli Kallio. *Esimiehen muuttuvat identiteetit: Narratiivinen tutkimus keskijohdon identiteeteistä ja samastumisesta organisaatiomurroksessa.* 2014.
55. Sirpa Tokola-Kemppi. *Psykoanalyyttisen psykoterapian merkityksiä kirjailijahaastattelujen valossa.* 2014.
56. Dhuana Affleck. *How does Dialogical Self Theory appear in the light of Cognitive Analytic Therapy? Two approaches to the self.* 2014.
57. Teemu Ratinen. *Torjuttu Jumalan lahja. Yksilön kamppailu häpeällistä seksuaalisuutta vastaan.* 2014.
58. Päivi Löfman. *Tapaustutkimus itseohjautuvuudesta sairaanhoitajakoulutuksen eri vaiheissa.* 2014.
59. Päivi Kujamäki. *Yhteisenä tavoitteena opetuksen eheyttäminen. Osallistava toimintatutkimus luokanopettajille.* 2014.
60. Henriikka Vartiainen. *Principles for Design-Oriented Pedagogy for Learning from and with Museum Objects.* 2014.
61. Päivi Kaakkunen. *Lukudiplomin avulla lukemaan houkutteleminen yläkoulussa. Lukudiplomin kehittämistutkimus perusopetuksen vuosiluokilla 7–9.* 2014.
62. Jari Kukkonen. *Scaffolding inquiry in science education by means of computer supported collaborative learning : pupils' and teacher students' experiences.* 2015.
63. Petteri Nieminen. *Unified theory of creationism — Argumentation, experiential thinking and emerging doctrine.* 2015.
64. Esa Antikainen. *"Me olemme postipaketteja": tutkimus Helsingin vammaiskuljetuksista vuosina 2002-2007.* 2015.
65. Anna Logrén. *Taiteilijapuheen moniäänisyys. Tutkimus mediavälitteisen ja (kuva) taiteilijalähtöisen taiteilijapuheen muotoutumisesta.* 2015.
66. Päivi Virkki. *Varhaiskasvatus toimijuuden ja osallisuuden edistäjänä.* 2015.
67. Terhi Nevalainen. *Pinkit piikkikorot. Chick lit -kirjallisuuden postfeministiset sisällöt ja lukijat niiden merkityksellistäjinä.* 2015.

68. Soile Tikkanen. *Development of an empathic stance. Dialogical sequence analysis (DSA) of a single case during clinical child neurological assessment procedures.* 2015.
69. Erja Laakkonen. *Aikuisuuden kulttuuriset kuvat 25-35-vuotiaiden naisten haastattelupuheissa ja naistenlehtiteksteissä.* 2015.
70. Wu Enqin. *What Does Design Expose? A Comparative Study of Finnish Modern Furniture and Chinese Ming-style Furniture.* 2015.
71. Vuokko Malinen. *Uuspari. Uusperheen parisuhde ja koulutuksellisen intervention mahdollisuudet sen tukemiseen.* 2015.
72. Juhana Venäläinen. *Yhteisen talous: tutkimus jälkiteollisen kapitalismin kulttuurisesta sommittumasta.* 2015.
73. Johanna Hokkanen. *Kotien lääkekasvatus.* 2015.
74. Eveline Omagano Anyolo. *Implementing Education for Sustainable Development in Namibia: A Case of Three Senior Secondary Schools.* 2015.
75. Satu Tuomainen. *Recognition and Student Perceptions of Non-formal and Informal Learning of English for Specific Purposes in a University Context.* 2015.
76. Jari Ruotsalainen. *Iholle kaiverrettu. Tatuoimisen kulttuurinen murros Suomessa.* 2015.
77. Sanna Hillberg. *Relativization in Scottish Standard English: a corpus-based study on newspaper language.* 2015.
78. Virpi Turunen. *Pellavalangan neulonta kotineulekoneella.* 2015.
79. Emilia Valkonen: *"Me myymme ja markkinoimme kursseja" Markkinaorientaation piirteet kansalaisopistoissa.* 2015.
80. Petri Haapa*, Suomalainen peruskoulun rehtori koulun tietokonepohjaisen hallintoohjelman käyttäjänä.* 2016.
81. Helena Valkeapää. *Maahanmuuttaja-afgaanien kotitalouden toiminta ja akkulturaatio.* 2016.
82. Airi Hakkarainen. *Matematiikan ja lukemisen vaikeuksien yhteys toisen asteen koulutuspolkuun ja jatko-opintoihin tai työelämään sijoittumiseen.* 2016.
83. Maria Kok. *Varjon kieliopillistuminen: itse-sanan paradigman rakenne ja merkityksenkehitys itäisessä itämerensuomessa.* 2016.
84. Päivi Vesala. *Koulupihan merkitys alakoulun oppilaille. Lapsen ja ympäristön vastavuoroisen suhteen analyysi.* 2016.
85. Izabela Czerniak. *Anglo-Scandinavian Language Contacts and Word Order Change in Early English.* 2016.

IZABELA CZERNIAK

Anglo-Scandinavian Language Contacts and Word Order Change in Early English *examines a major linguistic change that took place during the Old and Middle English periods and resulted in the establishment of a relatively rigid SVO word order. Among the various factors promoting this change, language contacts with the early Scandinavian population have often been mentioned as providing an important early input. This study investigates the extent of this influence and assesses its significance.*

UNIVERSITY OF
EASTERN FINLAND

uef.fi

PUBLICATIONS OF
THE UNIVERSITY OF EASTERN FINLAND
Dissertations in Education, Humanities, and Theology

ISBN 978-952-61-2114-7
ISSN 1798-5625