Terttu Nevalainen*, Tanja Säily, Turo Vartiainen, Aatu Liimatta
and Jefrey Lijffijt

History of English as punctuated equilibria? A meta-analysis of the rate of linguistic change in Middle English

Abstract: In this paper, we explore the rate of language change in the history of English. Our main focus is on detecting periods of accelerated change in Middle English (1150–1500), but we also compare the Middle English data with the Early Modern period (1500–1700) in order to establish a longer diachrony for the pace at which English has changed over time. Our study is based on a meta-analysis of existing corpus research, which is made available through a new linguistic resource, the Language Change Database (LCD). By aggregating the rates of 44 individual changes, we provide a critical assessment of how well the theory of punctuated equilibria (Dixon, Robert M. W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press) fits with our results. More specifically, by comparing the rate of language change with major language-external events, such as the Norman Conquest and the Black Death, we provide the first corpus-based meta-analysis of whether these events, which had significant societal consequences, also had an impact on the rate of language change. Our results indicate that major changes in the rate of linguistic change in the late medieval period could indeed be connected to the social and cultural after-effects of the Norman Conquest. We also make a methodological contribution to the field of English historical linguistics: by re-using data from existing research, linguists can start to ask new, fundamental questions about the ways in which language change progresses.

*Corresponding author: Terttu Nevalainen, University of Helsinki, Helsinki, Finland, E-mail: terttu.nevalainen@helsinki.fi. https://orcid.org/0000-0003-3088-4903
Tanja Säily, Turo Vartiainen and Aatu Liimatta: University of Helsinki, Helsinki, Finland, E-mail: tanja.saily@helsinki.fi (T. Säily), turo.vartiainen@helsinki.fi (T. Vartiainen), aatu.liimatta@helsinki.fi (A. Liimatta). https://orcid.org/0000-0003-4407-8929 (T. Säily). https://orcid.org/0000-0002-4760-750X (T. Vartiainen). https://orcid.org/0000-0001-9056-1087 (A. Liimatta)
Jefrey Lijffijt: Ghent University, Ghent, Belgium, E-mail: jefrey.lijffijt@ugent.be. https://orcid.org/0000-0002-2930-5057

1 Introduction

The broad outline of language contact in the British Isles is well known, and the English language itself has quite a long documented history. Much of this textual evidence has been digitized, making it possible to trace changes in the language empirically over some 12 centuries. As a result, a fair amount of quantitative, corpus-based work has accumulated in the last 30 years, which in turn makes it possible to ask to what extent these findings might converge over time. In other words, besides traditional research summaries, a historical linguist can now use this data accumulation to carry out meta-analyses of a number of studies to ask new questions about linguistic change.

Meta-analyses are routinely performed in fields such as medicine to find a *common effect* in quantitative evidence drawn from earlier studies (e.g. Cooper et al. 2009). Such analyses continue to be rare in most branches of linguistics with the notable exception of applied linguistics. For example, the articles included in Norris and Ortega (2006) meta-analysed quantitative studies on topics ranging from the effectiveness of corrective feedback to adult second-language learners' access to Universal Grammar. Durrant (2014) used meta-analysis to determine the correlation between learner knowledge of collocations and their frequency in corpora across 19 previously reported tests, and Goo et al. (2015) carried out a meta-analysis of 34 studies on explicit and implicit instruction in L2 learning.

With the body of quantitative research available on linguistic processes in earlier English, the common effect that we will focus on in this study is the rate of change over time. Our working hypothesis is informed by the notion of *punctuated equilibrium*. It is uncontested that events such as the Germanic invasions of Britain in the fifth century and the Norman Conquest in the 11th century had far-reaching consequences for the shape of the English language. Following Dixon's punctuated equilibrium model (1997), such events could be regarded as major external punctuations in the development of the language, followed by periods of relative stability (Bergs 2005: 53–55). Bergs goes as far as to argue that it is possible to agree on some average rates of change for the language system as a whole. Any "minor" punctuations in the development of the language are naturally also of interest to the historical linguist. One that has been proposed using data from the *Corpus of Early English Correspondence* is the effect of the English Civil War, which appears to have accelerated certain linguistic changes in progress in mid-to-late 17th century

following the vastly increased geographical mobility and contact brought about by the war (Raumolin-Brunberg 1998).

Both the major and minor punctuation models proceed from society to language. We would like to propose a complementary approach by proceeding from language to society. In this approach, we perform a systematic corpus-based study of the varying rates of a large number of linguistic changes, and compare the rates of change with language-external events over time. This undertaking has become possible with the Language Change Database (LCD), a new online resource, which currently draws together the results of c. 300 corpus-based studies on the history of English and which in the future can be regularly updated by members of the research community. We envisage that the LCD can be used to carry out meta-analyses of a large number of linguistic processes with the aim of discovering whether and to what extent their rates of change – acceleration or deceleration – converge on certain external events or periods of time (Nevalainen et al. 2016).

In this exploratory paper, based on the standardized numerical data included in the LCD, we will carry out a meta-analysis of 44 linguistic features undergoing change in the late medieval period. We will introduce the major candidates for punctuating events at the time, including the Norman Conquest and the Black Death (bubonic plague). To be able to assess our findings in terms of the stasis vs. punctuation hypothesis, we will critically review the data sources on which they are based and the ways in which these sources might have been affected by external upheavals. As regards language change, it is usually the case that written language follows spoken language developments with some temporal delay. Assuming that most changes are indeed initiated in the spoken medium, we should probably not expect to find any effects in the written record immediately after the punctuating event.[1]

The paper is organized as follows: by way of background, Section 2 is a brief overview of the early history of the English language and language contact in Britain focusing on the Norman Conquest. Section 3.1 introduces the punctuated equilibrium model and Section 3.2 discusses the general applicability of the model to the history of English and to the S-curve model of linguistic change; against this backdrop, Section 3.3 outlines the working hypotheses of our study. Section 4 introduces our materials: the LCD and the source material of the studies examined, i.e. the Middle English part of the *Helsinki Corpus of English Texts*. Section 5

---

1 Analysing the distance between written and spoken texts has received a good deal of attention over the years; for the history of English, see e.g. Samuels (1972: 4–8, 109–111); Kytö and Rissanen (1983), and more generally, e.g. Koch and Oesterreicher (1985); Biber and Conrad (2009).

discusses two possible methods of accounting for the rate of linguistic change. Our findings are presented and analysed in Section 6. We discuss these findings more extensively in Section 7, and reach our conclusions in Section 8. It is shown that, reflecting certain social and cultural developments catalyzed by the Norman Conquest, the textual record we studied points to one major juncture of accelerated linguistic change in Middle English.

2 Background

2.1 Early history of English

Language contact has a long history in the British Isles. The earliest inhabitants whose language is known to us were speakers of insular Celtic languages. During the Roman Rule of southern Britain (from 43 BCE to about 410 CE), varieties of Celtic co-existed with Latin, which was then primarily spoken and written by the Roman elite but persisted in various functions in the British Isles throughout the medieval period.[2] Beginning in the fifth century, Germanic people – Angles, Saxons, Jutes and Frisians – from the North Sea coast started raids around the coast of England. With time they colonized much of the territory between them, and their language eventually replaced the native Celtic in these areas (Filppula and Klemola 2009; Higham and Ryan 2013). The extent to which insular Celtic had a role to play in the overall structural simplification of English is a matter for debate (Trudgill 2016: 325–329).

From the late eighth century on, Britain became the target of raids by Vikings from areas of (modern) Denmark, Norway and Sweden as part of their military and mercantile expansion. As a result, Danes conquered East Anglia and large parts of Mercia and of the north, an area known as the Danelaw, and Old Norse (also called Scandinavian) was spoken in this area. The Anglo-Saxon speaking kingdom of Wessex continued to rule in the western part of Mercia and in the south (Figure 1).

It has been argued that, via koineization, the situation gave rise to "the amalgamation of English and Norse", which took place at different times in different places (Warner 2017: 385; Townend 2012). The stronger – and widely contested – claim based largely on syntax has also been made that Middle English could in fact be classified as Anglicized Norse (Emonds and Faarlund 2014).

---

2 For a map of the Roman empire at its largest, see Leiwo, in this issue. Overviews of the history of Britain in relation to the English language are available in a wide range of textbooks and reference works such as Crystal (2003); Bergs and Brinton (2012); Mugglestone (2012); Nevalainen and Traugott (2012), as well as a number of well-documented Wikipedia articles.
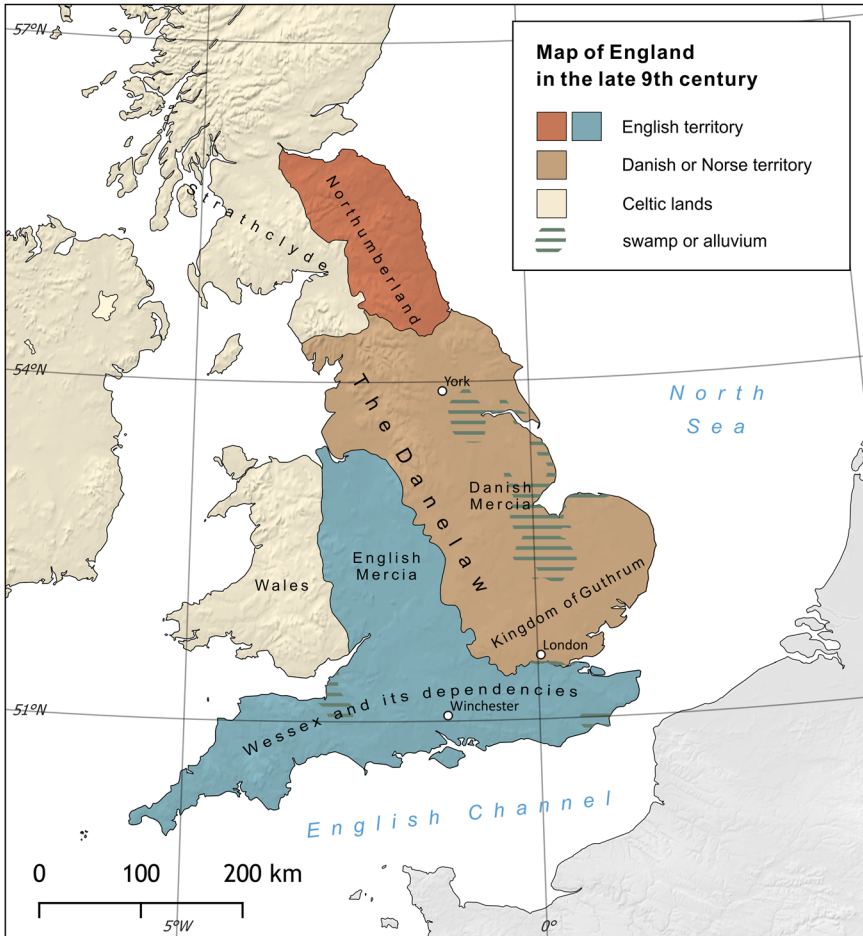
Figure 1: Map of England in the late ninth century, showing the extent of the Danelaw.[3]

2.2 The Norman Conquest and its aftermath

In northwestern Europe, Viking operations also extended to the coasts of what are now France and the Netherlands. In the 10th century, they culminated in the conquest and settlement of an area that corresponds to present-day Normandy, named after its inhabitants, the Normans. Apparently, Vikings integrated rapidly

3 Redrawn by Sakari Sarjakoski from https://en.wikipedia.org/wiki/Wessex based on the *Atlas of European History* by Earle W. Dowe, G. Bell and Sons, London, 1910 (England and Wales at the time of the Treaty of Chippenham AD 878).
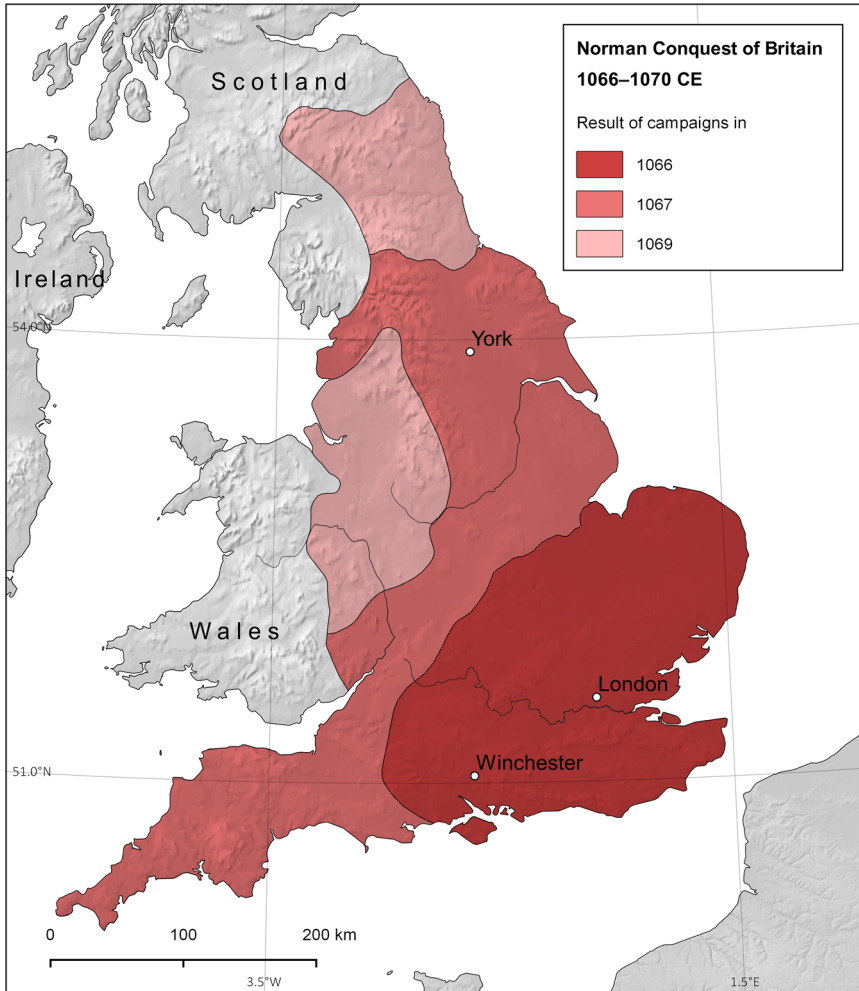
Figure 2: A map of Britain during the Norman Conquest.[4]

into the local community partly due to intermarriage, and their language declined, leaving only a number of Norse words in the local Romance dialect, called Norman French and later, in England, Anglo-Norman (Renaud 2008: 456–457).

4 Redrawn by Sakari Sarjakoski from https://etc.usf.edu/maps/pages/6800/6860/6860.htm based on Plate XXXI in *New Historical Atlas and General History* by Robert H. Labberton, Townsend MacCoun, New York, 1886.

In the 11th century, there was a succession crisis in England which came to a head in 1066 with several claimants to the English crown. In September, one of the prominent contenders, William, the Duke of Normandy, defeated Harold Godwinson, the last Anglo-Saxon king of England, at the battle of Hastings. William proclaimed himself king of England and invaded the rest of the country (Figure 2).

Bartlett (2000: 1) describes the aftermath of the Norman Conquest: "England in 1075 was a conquered country. Many of those who had recently been its rulers, the Anglo-Saxon landed class, were dead, exiled, or pressed down into the ranks of the peasantry. A small armed group speaking a language incomprehensible to the majority of the population controlled virtually all landed wealth."

The landed elite was gradually nativized in the course of the 12th century, and English–French bilingualism became the norm among the higher, literate social ranks but also spread further down on the social scale to professional and trade communities (Schendl 2012; Timofeeva and Ingham 2018).

The dynastic link between England and France was severed at the beginning of the 13th century during the reign of King John, who lost control of the Duchy of Normandy and most of his French holdings following the conflict with King Philip II of France. This defeat gave rise to subsequent conflicts between English and French rulers, notably the Hundred Years War (1337–1453), in which English monarchs laid claim to the French throne on grounds of their Norman ancestry and past possessions.[5]

From the 12th to the 14th century, English, Latin and French were used for different social and intellectual functions (Strohm 2006). In post-Conquest England, Anglo-Saxon was first replaced by Latin in the public domain, but French, both continental and insular, soon gained functions as a documentary and literary language alongside Latin, a situation that prevailed in a few domains such as law and religion until the end of the Middle Ages (Bartlett 2000: 482–524).

Social, political and demographic factors contributed to the gradual decline of French and rise of English as a written language in the 14th century. Schendl (2012: 507–508) observes that this development was furthered by the Hundred Years War and the mid-century plague epidemics, which reduced the population of England at least by one third and led to a severe shortage of labour, thus indirectly promoting the prestige of English. Growing urbanization and the rise of an English-speaking merchant class were further factors in the process.

The Black Death also hastened the decline of spoken Anglo-French: the pandemic had a particularly high incidence among the clergy, monasteries and convents that had been responsible, among other things, for providing English-speaking children with early instruction in spoken French, which had served as a

---

5 For further information, see Thomas (2008) and https://en.wikipedia.org/wiki/Hundred_Years %27_War, and the references therein (accessed 22 November 2018).

vehicular language for learning Latin in grammar schools (Ingham 2014: 640–641, 2017). In retrospect, the single most significant effect of late medieval multilingualism was the influence that French had on English lexis and phraseology (Durkin 2014). We will return to this topic in Section 7.2.1.

3 Modelling the rate of linguistic change

3.1 Dixon's model

The punctuated equilibrium model was first presented by Niles Eldredge and Stephen Jay Gould in 1972 in their discussion of evolutionary change in biology. Dixon (1997) transferred the model to the context of language change and applied it to the prehistory of some Australian indigenous languages. Dixon (1997: 67) suggests that language history typically comprises long periods of stability during which languages have coexisted in a given region without any major changes taking place. But from time to time this equilibrium can be punctuated by an event which causes abrupt changes in the linguistic situation.

According to Dixon (1997: 67–85), most examples of such punctuations stem from non-linguistic factors, and are due to various natural causes, as for example floods, drought, and the fall or rise of sea-levels. They can also be the consequence of material innovations, including the invention of agriculture, of new weapons or implements and new means of transportation. Moreover, the development of aggressive tendencies in a population – conquests, religious and political expansionism, new social hierarchies – can trigger sweeping changes in the linguistic situation. Writing and other forms of communication can serve as concomitant factors in these processes of punctuation (Dixon 1997: 75). Another major source for punctuations in the model comprises various geographical parameters, including invasions of and expansion into previously occupied territory, and population expansion into uninhabited territory, which can lead to splits of political groups and of languages. These can of course combine with any of the developments mentioned above, such as material innovations and religious and political expansionism.

3.2 Approaches to the rate of linguistic change in earlier English

Our approach to the punctuated equilibrium model is informed by the commonly held view articulated by Denison (2003: 68) "that Old English and (late) Modern English are relatively invariant, whereas Middle (and possibly early Modern)
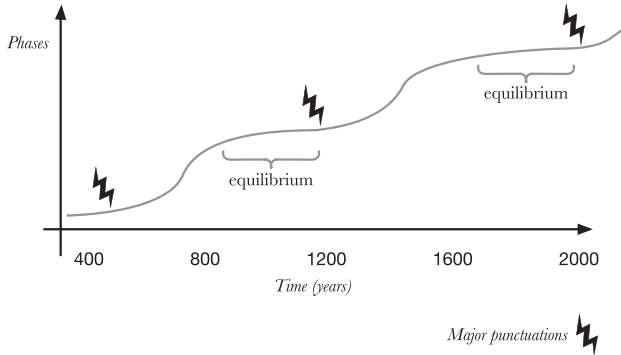
Figure 3: History of English as punctuated equilibria (modified from Bergs 2005: 54).

English show rapid change of all kinds". Denison further considers the idea that the fifteen-hundred-year history of the English language might be thought of as one big S-curve of change. This is also implied by Figure 3, modified from Bergs (2005: 54), but with the difference that Bergs marks periods of relative stability punctuated by periods of rapid change and associates these periods with language-external events of different kinds.

In principle accommodating this type of fluctuation, Denison (2003: 67) introduces an analytical aspect into his S-curve model in that he suggests "superposing a number of graphs to see where they bunch in historical time, rather like looking at isoglosses in the hope of spotting a significant dialect boundary. And rather than the S-curves themselves, peak values of the *rate* of change (first differential of the S-curve) would have to be plotted to see whether they tend to coincide". Clearly the rate of change can be approached at different levels of generalization and with different methods. Denison's focus is on the rate of change of individual processes and their contribution to the overall shape of linguistic change over time. This is also the approach we adopt in our meta-analysis: rather than charting the totality of changes in the language in a given period qualitatively, we analyse the pace of individual processes quantitatively to see "where they bunch in historical time".[6] A more schematic approach is presented by Aitchison (1981: 100), who visualizes the progress of an individual process of change in its different linguistic contexts by drawing a cumulative S-shaped curve

---

6 We would like to thank an external reviewer for drawing our attention to this distinction. Ideally, a fully comprehensive account of the pace of change in a language would account for the rate at which all linguistic changes, identified as such, progress in real time.

of change as a series of successive overlapping smaller S-curves (Nevalainen and Raumolin-Brunberg 2017: 54–55).

Analysing individual changes in progress, Kroch (1989) argues that patterns that map a process of change in its different linguistic environments follow the Constant Rate Hypothesis. However, he makes a difference between functional and stylistic variation – "contextual effects" – in the use of the incoming form, and the more abstract change in grammar, which "seems to proceed at the same rate in all [linguistic] contexts" (Kroch 1989: 199). Kauhanen and Walkden (2018: 285) describe the usual method for detecting this Constant Rate Effect (CRE) by fitting a logistic curve to each linguistic context separately. If variation among the time-independent rate-of-change or "slope" parameters for the curves falls within what is considered a reasonable confidence interval, the change is taken to proceed at the same rate in all contexts. On the other hand, variation is allowed among the "intercept" parameters, which mark the points of greatest growth along the time axis. The authors reject this standard operationalization of CREs partly because it leaves variation among the external, intercept parameters unexplained (Kauhanen and Walkden 2018: 286). As Kauhanen and Walkden (2018: 287) note, contextual effects can relate to between-speaker effects or even contingencies, which are not addressed by the standard CRE model. Our interest in this article is in the observed text frequencies of linguistic features and any changes in the rates at which they progress over time. Hence our approach does not abstract away from this variation, whatever its external source might be, and our quantitative methods are constructed accordingly (Section 5).

3.3 Working hypotheses of the study

Focusing on the later medieval period, we consider several external, non-linguistic events as possible candidates for punctuations in the rate of linguistic change in Middle English. The first one is the Norman Conquest in 1066. Incidentally to it, the administrative, political and cultural activity was transferred from Wessex to the East Midlands with London as the capital city. Another major punctuating event that we consider is the Black Death. Our interest in this paper lies in the potential effects of these events on the rates at which processes of language change can be observed to unfold over time by studying the textual record that has come down to us.

As suggested by Dixon (1997) and Bergs (2005), the Norman invasion caused one of the major punctuations in the history of the English language in that it made the country trilingual, with French and Latin taking over the functions of English as a literary medium. Discussing the state of linguistic equilibrium in an area being

punctuated by invasion, Dixon (1997: 84–85) singles out the Norman Conquest as a special case of the invaders' language eventually falling into disuse. He observes that this could happen in the special circumstances where the invaders were relatively few in number and came from a very similar culture to that of the English at a time when there was only one major language spoken in England (Dixon 1997: 85; Thomason 2008: 48). This superstrate shift left English with a large number of French loanwords but the extent to which it had a direct impact on the rate of ongoing processes of change in English remains an empirical issue.

Another source of disruptive influence, following from the Norman invasion, is geographical, and it involves the transfer of the country's power base from Wessex to the East Midlands and London, the largest city in the country (Keene 2000: 99–101). Subsequent to the Norman Conquest, London was made the seat of royal administration, and by 1200 the Palace of Westminster was established as the principal royal residence and permanent administrative centre of the kingdom. The position of London was strengthened by its increasing volume of trade and the agricultural resources in the surrounding areas capable of supporting growing urbanization (Britnell 2006: 150–152). These developments contributed to the political, economic and cultural dominance of the capital region, and transferred much of the country's earlier documentary and literary activity away from the south-west and West Midlands, which meant that when the English language was eventually re-established in these functions from the mid-14th century onwards, the dominant written variety had shifted to the East Midlands, an area which had been linguistically influenced by earlier Scandinavian contacts (see Figure 1). We shall call these processes collectively Post-Invasion Effects.

A further source of punctuation to consider in the late medieval period is the impact of the Black Death on the population of England. This impact was aggravated by a series of bad harvests, famine, and a subsequent low replacement rate of lost population accompanied by labour shortage in the decades following the first severe outbreaks of the plague in 1348–1349. Their joint impact is estimated to have brought the English population down at least by one third or, according to some estimates, by as much as 50 per cent, from about four million down to two million (Keen 1990: 27–47; Hatcher 1994: 71; Slack 2012: 35–41). In comparison, population size was only marginally affected by the intermittent campaigns of the Hundred Years War. It has been estimated that, except for the upper social ranks, only a tiny fraction of the population ever saw war service during this period (Keen 1990: 134–140).

The impact of population size on the rate of language change remains an empirical question (see Sinnemäki 2020, this issue). Topics like word loss and population size have been investigated, but the results have so far been inconclusive. Greenhill et al. (2018: 16) identified some significant correlations between

population size and the rate of word loss in the Indo-European languages they tested but could not establish similar correlations for Austronesian and Bantu languages. The situation in post-plague England presents a scenario for the study of potential correlations between a radical reduction in population size and the rate of language change. It is, however, complicated by the fact that population size alone cannot be separated from the other social consequences of the Black Death and the Hundred Years War that affected, for example, social network structures in the language community by vastly increasing social and geographic mobility (Keen 1990: 44, 140).

4 Material

4.1 Language Change Database

The data for the meta-analysis have been taken from the **Language Change Database (LCD).** The LCD is a new linguistic resource – a research database which draws together a significant collection of existing corpus-based research into the history of English. Each individual research article is represented by one annotated database entry in the LCD. Each entry contains basic bibliographical information, including the author of the article and the source publication, while the keyword-based annotations target e.g. the grammatical, dialectal and sociolinguistic phenomena studied in the article, as well as the corpora and other databases used in the study. In addition to these keywords, the entries include an abstract of the article and a concise summary of its main results. These are intended to provide the users of the database with a convenient way to assess whether or not the research presented in the article is relevant to their purposes. Importantly, the database entries also include numerical data as reported in the original articles. These data are stored as Excel files, which the end users of the LCD can download on their personal computers for further examination and reuse.

    The LCD is currently in beta stage, and the development team has thus far taken on the responsibility for updating it with new entries. In the future, however, other researchers will be given access to the backend of the database, and they will be able to supplement it with detailed information of their own research. At its present state, the LCD comprises information from c. 300 corpus-based research articles that study the history of the English language from Old English to the present day. In our view, the size of the database has now reached a point where it can be used as a baseline for new research; more specifically, we can now abstract away from the micro-level research questions studied in the original articles and

use the existing data to study some fundamental questions concerning the nature of language change from a macro-level perspective.

The LCD comes with a customized grammatical model, which is based both on the comprehensive grammatical descriptions found in the large reference grammars of Present-day English (Huddleston and Pullum 2002; Quirk et al. 1985) and on our extensive reading of earlier research on the history of English (see Nevalainen et al. 2016). This combinatory approach to building the grammar component was motivated by two reasons. First, the synchronic reference grammars of English quite obviously ignore those features of grammar that have become obsolete over time, which is why a more comprehensive view of the long diachrony of the language is required. Second, by consulting existing research on the history of English, we have been able to gain an understanding of what the research foci in our field are and what kind of terminology scholars have commonly used to describe the phenomena studied. In sum, the grammar component of the LCD is designed to provide the users of the database with an interface that would most conveniently suit their purposes, while being theoretically as neutral as possible.

The architecture of the LCD allows the users to query the database according to varying degrees of granularity (Figure 4). For instance, the LCD's grammar component is hierarchically structured, and all grammatical keywords with which the database entries have been annotated are automatically included in their related superordinate categories. For example, an entry that is annotated for

Figure 4: The search interface of the LCD.

modal verbs will also be found by a query targeting **verbs**, and an entry annotated for **personal pronouns** will be found when articles that study **pronouns** are searched in the database. The hierarchical structure also ensures that the end users will be able to find the information in which they are interested with high precision, which reduces the need for manual post-editing of the query results. In addition to the grammar component, the LCD includes detailed information about the composition of various corpora, which can be used, for instance, when searching for research on specific genres (e.g. according to the detailed genre classification of the *Helsinki Corpus of English Texts*). Once published, the LCD will be accompanied by a tool designed to streamline the workflow related to meta-analysis, which further reduces the need for manual data processing (LADA; see Kesäniemi et al. 2018). The LADA tool can access the corpus information included in the LCD, which makes it useful for a variety of purposes in the meta-analytical process, such as the normalization of raw frequencies according to specific time periods or genres.

All the data for the meta-analysis presented in Section 6 have been taken from the LCD. We filtered the data according to three criteria:

 i. corpus (only articles that made use of the *Helsinki Corpus* were included),
 ii. word classes (only articles that studied nouns, pronouns, adjectives, determiners and verbs were included), and
iii. time period (only articles that studied Middle English were included; these studies potentially extend the investigation to Old English or Early Modern English)

We considered the first criterion to be important because of potential issues related to the comparability of results across different corpora. The second criterion was motivated by the insight that different levels of linguistic organization may change in different ways and at different rates (see Lass 1997). It would therefore be sensible to aim at a sufficiently representative sample of language. Finally, the Middle English period provides us with an opportunity to study the potential correlation of several language-external events and the rate of linguistic change.

In all, we were able to locate 44 individual linguistic changes that matched our criteria (for more details, see the list of primary sources under References). These changes represent a wide variety of grammatical phenomena, including the development of modal auxiliaries, degree words, indefinite pronouns, quantifiers, relative adverbs and derivational morphemes (the nominal suffixes *-ness* and *-ity*). The frequency data related to the variation and change of these phenomena will serve as the baseline for the meta-analysis discussed in Section 6 below. Although we had filtered the data according to the three criteria mentioned above, the selection process was still completely random in the sense that we had not screened the data beforehand. Indeed, the results that we obtained did not quite

match our initial expectations, which can perhaps be taken to support the idea that data selection was indeed carried out in an unbiased way.

4.2 *Helsinki Corpus*

The *Helsinki Corpus of English Texts* (HC 1991), on which all of the studies included in this meta-analysis are based, is a diachronic corpus of c. 1.5 million words, spanning from Old English to Early Modern English. In order to establish how compatible and comparable the data are between the different subperiods of the corpus, and to ensure that any simultaneous changes we may find are not merely a corpus artefact, it is important to understand the structure and contents of the corpus.

As discussed above, the main focus of our meta-analysis will be on Middle English. In the *Helsinki Corpus*, the Middle English period has been divided into four subperiods: ME1 (1150–1250), ME2 (1250–1350), ME3 (1350–1420) and ME4 (1420–1500). The corpus contains a representative sample of written English language from each of the periods, although the distribution of word counts is uneven because of the historical circumstances related to the Norman Conquest and the diminishing role of English as a written language; this is even more noticeable in the last Old English period in the corpus, 1050–1150 (see Table 1; Kahlas-Tarkka et al. 1993). Moreover, the proportion of texts which are re-productions of earlier manuscripts varies between the periods. Most texts in Early Middle English (ME1–ME2) are based on a historically prior text, while in Late Middle English (ME3–ME4) the majority of the texts are contemporaneous (Nevanlinna et al. 1993). Therefore, the earlier periods may appear more conservative than the later ones because of differences in the historical continuity of text production.

Table 1: Word counts of the periods OE4–ME4 in the *Helsinki Corpus*.

Old English	Timeline	Word count
IV	1050–1150	67,380

Middle English	Timeline	Word count
I	1150–1250	113,010
II	1250–1350	97,480
III	1350–1420	184,230
IV	1420–1500	213,850

Because the cultural and literary context changed in the course of the Middle English period, with different English dialects dominating at different times and with new written genres emerging and old ones transformed, the exact dialect and genre makeup of the corpus must by necessity also change from one period to the next. In order to gauge whether variation in the rates of change between time periods might be better explained by the changing genre and dialect composition of the corpus rather than by the linguistic and extra-linguistic processes described above, we graphed the breakdown of the corpus word count by dialect, text category, and mode of production (Figures 5–7). For additional context, Figures 6 and 7 extend these characterizations to the Early Modern English period. We found considerable inter-period continuity but also substantial differences between the periods.

In Figure 5, we can see that the Midlands dialects account for over three quarters of the words in the corpus in each subperiod, with the East Midlands dialect leading in all but the first period. In the second half of the Middle English period, the Northern dialect emerges and the proportion of the Southern dialect decreases, but the change is small overall.

The text category breakdown in Figure 6 is based on the prototypical text categories included in the *Helsinki Corpus*, but we further merged some of the categories to create a smaller number of more general supercategories. The figure shows that narrative and religious texts command the majority in the corpus throughout the Middle English period, with new genres starting to appear towards the end of the period. The mode of production (Figure 7) is also relatively balanced, with the obvious exception of ME2, which was the low point in English vernacular text production (see further Nevanlinna et al. 1993).
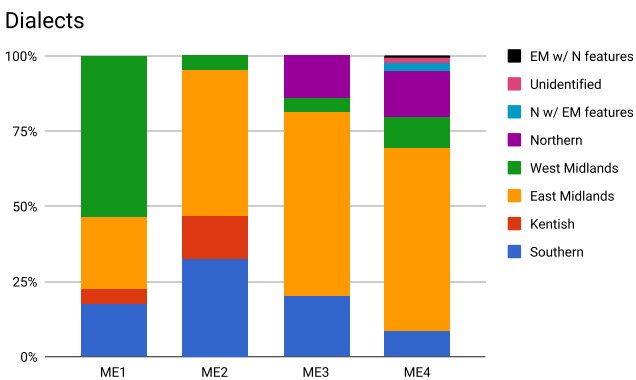

Figure 5: The dialect distribution in the Middle English section of the *Helsinki Corpus*.
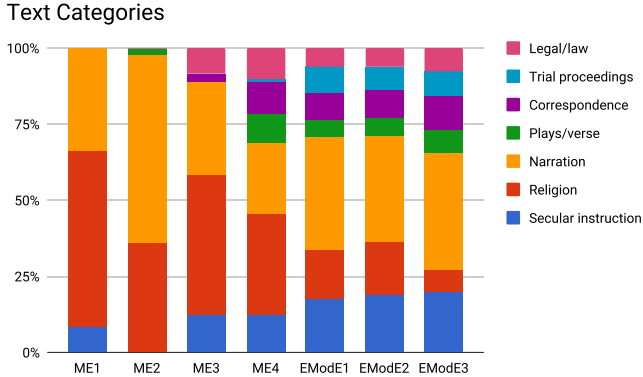
Text Categories

Figure 6: The distribution of text categories in the ME and EModE sections of the *Helsinki Corpus*.

5 Methods for analysing the rate of change

Language change is not a stable process, and as argued above, the rate at which a language changes is itself subject to change over time. A question that arises is whether we can quantify the rate at which a language changes in a meaningful manner. How to quantify the rate of change may depend on the type of data under consideration. Here the analysed data consist of raw or normalized text frequencies which are aggregated into bins that correspond to time periods.

A direct approach to analysing the rate of change would be to look at the absolute differences between time periods for one or more linguistic variables.
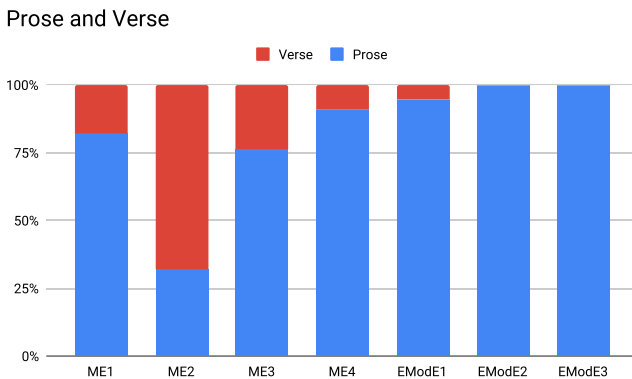
Prose and Verse

Figure 7: The distribution of modes of production in the ME and EModE sections of the *Helsinki Corpus*.

However, this approach is problematic from a statistical perspective. More specifically, since the frequencies are based on collected samples, the question is whether the expected variability associated with the sampling process is associated with the magnitude of the frequencies of the linguistic features.

Arguably, we have no good model that gives us this variability. Several studies have looked at the dispersion of lexical items across a corpus in terms of their frequency as well as properties such as semantic type (e.g. Altmann et al. 2009; Lijffijt et al. 2016). From these studies, we already see that assessing the expected variability is a complicated matter: in general the variance increases with the frequency of a lexical item, yet more frequent words appear to be generally better dispersed, and the degree of dispersion also depends on the word type.

In summary, there are two things we know that are essential to designing a sensible measure for calculating the rate of change. (1) When we compare or aggregate several linguistic features, it is necessary to normalize their frequencies because the variance associated with sampling depends on the size of the corpus and the frequency of the feature. (2) Lacking a good model, it seems unjustified to use a parametric approach (e.g. assume independence between all instances, which would allow us to compute the expected variance using the binomial distribution). This means that we need to base the normalization directly on the frequencies.

It seems that there are at least two methods that can be adopted to meaningfully compare the observed changes in the frequencies of linguistic features over time. For each feature, we start by computing the difference in its frequency in one period against its frequency in the previous period. Note that if we start with four periods, this means that there are three such differences. In the first approach, which we call *relative change rate* (*deviation-normalized change rate*), we compute the standard deviation of these differences, and the final rates of change for each linguistic feature are given by dividing the differences by the standard deviation (Example 1).

1. Normalized frequency/10,000 words (example)

ME1	ME2	ME3	ME4
30	40	20	10

2. Difference between adjacent periods

ME1…ME2	ME2…ME3	ME3…ME4
40−30 = 10	20−40 = −20	10−20 = −10

3. Difference normalized by standard deviation = **rate of change** between the periods

ME1…ME2	ME2…ME3	ME3…ME4
10/15.2753 = **0.6547**	−20/15.2753 = **−1.3093**	−10/15.2753 = **−0.6547**

Example 1. Computing the rate of change, method 1.

Likewise, in the second approach, which we call *proportional change rate* (*difference sum normalized change rate*), we compute the differences as above, but then also take the absolute values. Notice that when a frequency for some period is lower than in the following period, the difference is negative. The absolute value is the same number but non-negative (zero or higher). We then compute the sum of these numbers and divide each absolute difference by the sum. One could multiply the resulting values by 100 to obtain the percentage of change between two periods (Example 2).

1. Normalized frequency/10,000 words (example)

ME1	ME2	ME3	ME4
30	40	20	10

2. Absolute difference between adjacent periods

ME1…ME2	ME2…ME3	ME3…ME4
|40−30| = 10	|20−40| = 20	|10−20| = 10

3. Proportion of overall change that occurs between the periods = **rate of change** between the periods

ME1…ME2	ME2…ME3	ME3…ME4
10/(10+20+10) = **0.25** (25%)	20/(10+20+10) = **0.5** (50%)	10/(10+20+10) = **0.25** (25%)

Example 2. Computing the rate of change, method 2.

Both approaches have their advantages. The first method is likely to be less easy to understand and hence the figures visualizing series of differences are more difficult to read. The second approach is both easy to understand and the scale is also

fixed (between zero and one), which additionally reduces the cognitive workload in reading figures that depict such series. However, in the second approach, we lose the directionality of the change. When we take the absolute value, we can no longer see whether the count is increasing or declining. Hence, during analysis we have to verify the direction of the change from the original series of differences. On the other hand, we are able to identify periods where the cumulative rate of change is at its greatest regardless of the direction of the individual changes.

In both cases, as the scale across measured phenomena becomes comparable, we may compare series of differences visually. However, care still needs to be taken both when making comparisons between phenomena, as well as when aggregating change rates over series of measurements that do not cover the same set of periods. For brevity, we take the approach of conducting any analysis first using all measured phenomena, and then consider the effect of removing the series that do not span the full period. As it turns out, none of the analyses are impacted by shorter series, and hence we leave the discussion of how to resolve this issue in a statistical manner for future work.

In mathematical terms, the methods can be described as follows.

- Given a series of n measurements
  $$x_1, x_2, \ldots, x_n$$
- Compute the $n - 1$ differences
  $$d_i = x_{i+1} - x_i \quad \text{for } i = 1, \ldots, n - 1$$
- Method 1 (relative change rate)
  $$r_i = \frac{d_i}{\sigma_d}, \quad \text{with } \widehat{\sigma}_d = \sqrt{\frac{\sum_{i=1}^{n-1} \left(d_i - \overline{d}\right)^2}{n-2}}$$
- Method 2 (proportional change rate)
  $$p_i = \frac{|d_i|}{\sum_{i=1}^{n-1} |d_i|}$$

6 Analysis of the rate of change in Middle English

As the first step of our meta-analysis, we normalized the frequencies of the 44 linguistic processes retrieved from the LCD and plotted all of the frequencies onto a line graph (Figure 8). However, this approach has some disadvantages: the overlapping lines are hard to distinguish, and the most frequent changes dominate the graph. Based on this graph, we are not able to say whether the rate of change is faster between some periods than others.

Our next step was to plot the frequencies on a logarithmic scale (Figure 9). While the figure is still quite messy, there do seem to be some interesting frequency
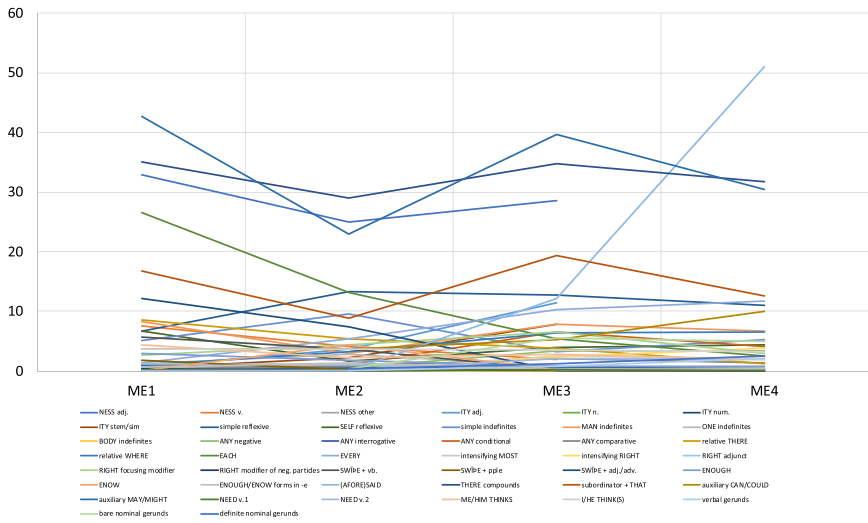
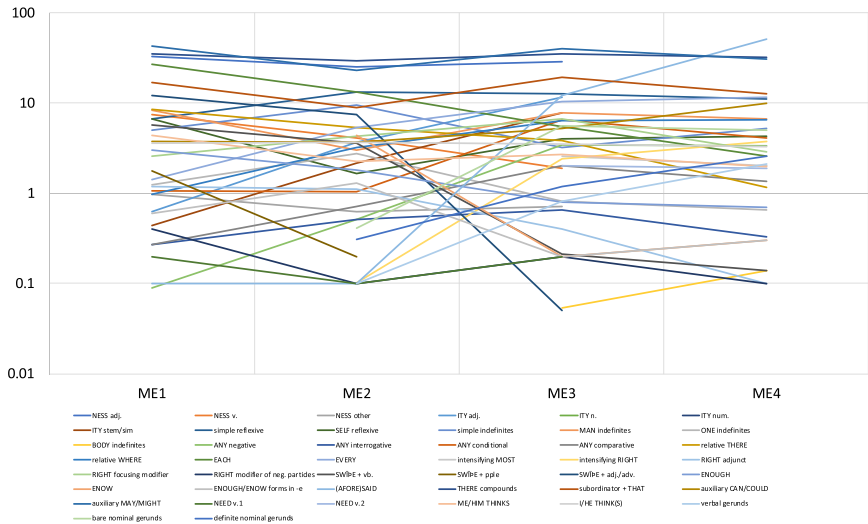**Figure 8:** Frequencies of 44 changes retrieved from the LCD, normalized to 10,000 words.

Figure 9: Frequencies of 44 changes (normalized to 10,000 words) on a logarithmic scale.

changes between ME2 and ME3 that cross over one or more orders of magnitude, implying a faster rate of change.

In order to quantify the rate of change across multiple changes more precisely, we used the second method described in Section 5 above, "proportional change rate". The results are illustrated in Figures 10–14. We first grouped the changes by type, as otherwise e.g. -*ness* and -*ity* would have skewed the results with seven different changes. This left us with six categories: derivational morphemes (-*ness*/-*ity*), gerunds, degree words, pronouns, clause-level phenomena (*there* compounds and subordinator + *that*) and verbs. Figure 10 shows the overall rate of change with this categorization. Instead of a line graph, we opted for a stacked graph that displays the cumulative effect of each category on the overall rate. In this visualization, the value of each category in a time period is added to the values of the categories below it on the graph, so that the height of the stack corresponds to the sum of the values in the time period. For convenience, we have further normalized these values (dividing them by the number of categories) so that when the sums for each time period are added together, their total sum is 1. This makes it easy to see that e.g. in Figure 10, the proportion of overall change that occurs between ME1–ME2 (i.e., the rate of change between those periods) is a little less than 0.3 or 30%, that between ME2–ME3 is somewhat below 0.5, and that between ME3–ME4 is somewhat above 0.2.

It is thus immediately obvious that the overall rate is the highest between the periods ME2 and ME3, and calculating the average change rates (as a complementary measure to the summed rates) confirms this finding. While this does not obtain for each individual change (e.g. verbs and gerunds show a slightly higher rate of change in ME3–ME4), the general trend is clear. As a sanity check, we have
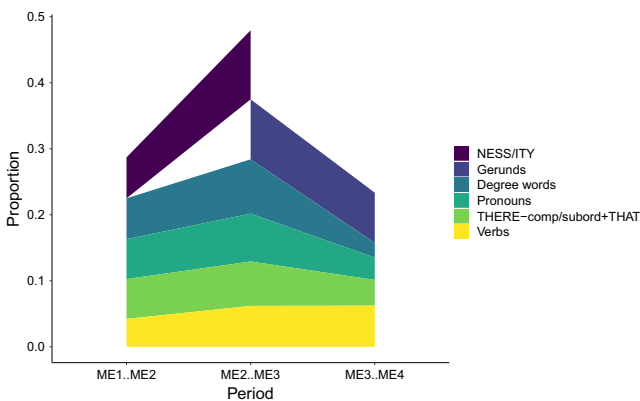

Figure 10: A stacked graph of the overall rate of 44 changes pooled into six categories. The areas in white indicate lack of data between the periods in question.

ordered the stack so that the phenomena with missing data and the greatest differences in the rate of change are on top, so that we can easily assess the effect of leaving them out. For instance, even if we exclude the derivational morphemes from the analysis, the cumulative rate is still the highest in ME2–ME3, as is the case if we leave out the gerunds and degree words.

As noted in the previous section, our method does not account for the directionality of the change. We therefore conducted a separate analysis of the 11 changes whose normalized frequency consistently increased over time (Figure 11). These incoming features provided the same result as above: the rate of change is the fastest in ME2–ME3. The rate in ME1–ME2, on the other hand, is surprisingly slow, considering that the compositional differences between ME1 and ME2 in the *Helsinki Corpus* seem to be at least as great as the differences between ME2 and ME3 (Section 4.2 above).

After looking at the incoming features, we turned to the six consistently recessive features (Figure 12). Here the results are quite different in that the rate of change is at its fastest between the first two periods, ME1–ME2, and decreases throughout the Middle English period. Note, however, that most of these features (e.g. *-ness*) are not outgoing in the sense that they would end up disappearing from the language completely.

While the focus of our analysis was on Middle English, we were also interested in seeing how the Middle English period compared to the Early Modern period in terms of rate of change. For this analysis we concentrated on changes for which we had data from ME1 to EModE3. Figure 13 shows that the overall rate of change peaks at ME2–ME3, providing further support for the significance of this period. Interestingly, for these changes that are ongoing throughout the Middle and Early Modern English periods, the rate of change tends to be greater in Middle English (ME1–ME3) than in Early Modern English.
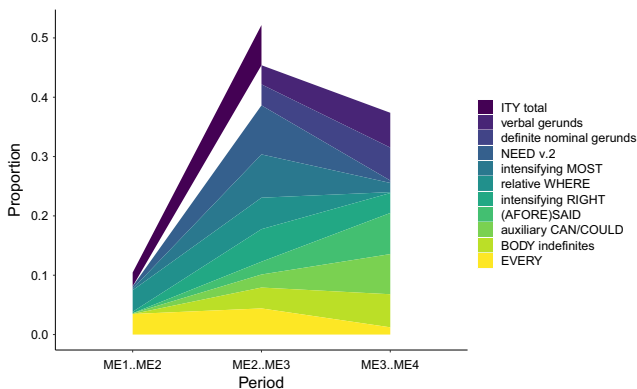

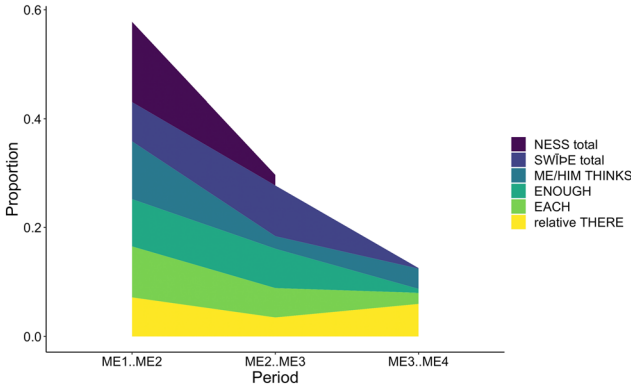**Figure 11:** Rate of change: incoming features.

Figure 12: Rate of change: recessive features.

As a final step, we decided to study the data by excluding the ME2 period from the analysis; as discussed in Section 4.2, this period stands out from the other Middle English periods both in terms of dialect distribution and the proportion of prose and verse texts. It is therefore possible that some of our results might arise as a consequence of the composition of the *Helsinki Corpus* instead of reflecting genuine differences in the rate of change across different periods. Figure 14 shows the rate of change from Middle to Early Modern English with ME2 excluded. Now the overall rate of change is at its highest between ME1 and ME3, and the Middle English period still stands out as more rapid than the Early Modern period,
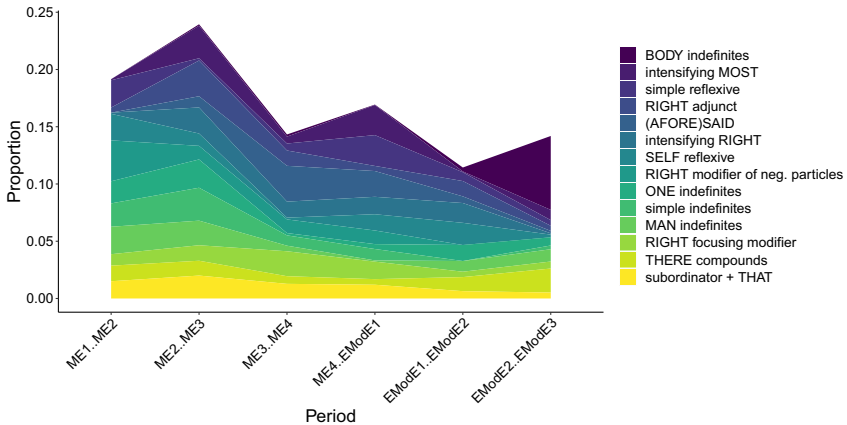

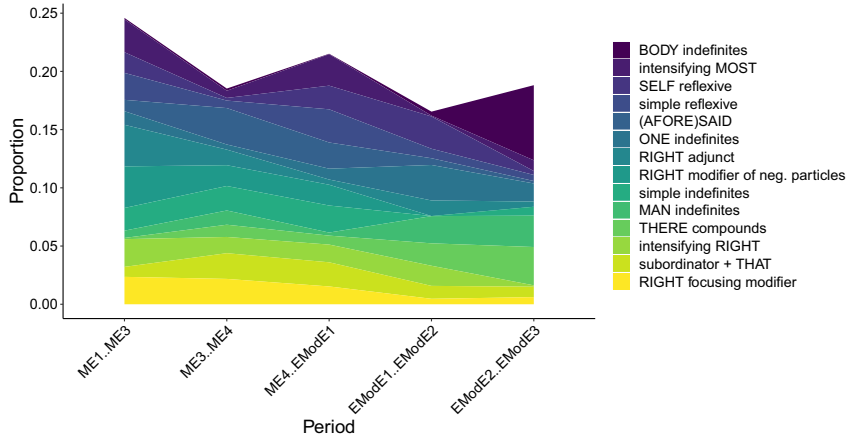**Figure 13:** Rate of change in the long diachrony.

Figure 14: Rate of change in the long diachrony, ME2 excluded.

especially if we disregard the *-body* indefinites, which only really start to increase in EModE2–EModE3.

7 Discussion

Any quantitative approach to language study is sensitive to the granularity of the data and of the linguistic analysis. Aggregate data comprising information from various genres, such as the material used in this study, shows the big picture over time but conceals genre, social group and register differences (see further e.g. Nevalainen 2015: § 4). On the other hand, raising the level of abstraction is necessary in studies of early English, where the time span covered is long and only small systematically compiled text corpora are available. Since our study is the first of its kind, our aim is simply to explore whether acceleration in the rates of documented processes of linguistic change could in principle be associated with abrupt external events and their impact on the speaker population.

7.1 Summary of results

In our study, we set out to explore the extent to which the rate of language change could be shown to be affected by language-external events that cause punctuations into a state of relative equilibrium and in this way accelerate change (Dixon 1997). To shed more light on this question, we carried out a meta-analysis of 44 processes of

language change that were ongoing in the Middle English period. Many of these processes had already started in Old English, and only one of the changes (*need* v.1, 'to compel') was completed before the Early Modern period. Thanks to the longevity of the grammatical constructions studied, we were not only able to examine the rate of language change within the Middle English period but also to compare the rate of change in Middle English to that in Early Modern English.

Our results show that the Middle English period was characterized by accelerated change in the long diachrony of English. This conclusion is in line with comments made in previous research, but to our knowledge this is the first time when the claim has been backed up by quantitative evidence from a large number of randomly selected changes. However, we also found that the rate of language change varied within the Middle English period. When recessive processes were examined independently, the observed changes were generally most pronounced between the ME1 and ME2 periods, while for incoming changes the rate of change peaked between ME2 and ME3. It should be noted, however, that the majority of the changes (27 out of 44) were neither unequivocally incoming nor recessive in the Middle English period. When these data were pooled together with the incoming and recessive changes, the period from ME2 to ME3 also emerged as the period with the fastest rate of change.

The recessive features that we examined represent a mixed bag of linguistic items, and they also change in different ways. Some items disappear from the language altogether (e.g. *swīþe*), while others show a steadier decline (*there*-relativizers; impersonal constructions such as *methinks*). Yet for others, the question is about the reorganization of grammatical functions (*enough*, *each*) or variation in text frequencies (*-ness*). For example, the decline of the quantifier *each* is in part explained by the fact that *every* started to be used in its stead in certain functions (e.g. to indicate non-individuated reference, as in *every man*), while the decline of *-ness* in Middle English is potentially related to the introduction and increased use of another nominal suffix, *-ity* (see Section 7.2.1 below).

Like the recessive features, the incoming features represent a variety of grammatical and lexical phenomena. It is interesting to note that while some of the incoming and recessive items can be regarded as variables (*-ness* and *-ity*, *where*-and *there*-relatives, *each* and *every*), their rates of change peak at different times. While there may be language-internal reasons to account for this disparity, the data are too varied to draw firm conclusions. In general, it is possible that lexical changes may proceed more rapidly than syntactic changes, and this may have an effect on some of our results; this is certainly something that we are interested in establishing in future research. However, there are also significant language-external events that may have affected the rate of language change witnessed in our meta-analysis, and we would like to conclude our paper with a discussion of

the potential effect of such punctuating events with particular focus on the Norman Conquest and the Black Death.

7.2 Punctuating events vis-à-vis our results

We now return to the three tentative hypotheses outlined in Section 3.3.

7.2.1 The impact of the Norman Conquest

The Norman Conquest made England trilingual. The extent to which French impacted on Middle English has been subject to debate over the years and ranges from the creolization hypothesis (Bailey and Maroldt 1977) to being limited to lexical borrowing. Although English had lost much of its inflectional morphology and came to have a fixed basic SVO word order by the end of the medieval period, the contention that Middle English was grammatically simplified and ultimately creolized as a result of the contact with French has since been abandoned. However, the impact of French on English lexis and phraseology is incontestable (e.g. Timofeeva and Ingham 2018). The revised *Oxford English Dictionary* (OED3) suggests that French occupied a 15% share of new words in English in the first 50 years after the Conquest, and that its share peaked at some 40% in the first half of the 14th century; if the mixed etymologies of French and Latin are taken into account, the proportion remained roughly at that level until the end of the 15th century (Durkin 2014: Figure 2.6). A fairly similar overall profile of the share of French of all loanwords is suggested by the estimates based on the *Middle English Dictionary* (Miller 2012: 161). These statistics suggest an increasing lexical input from French in Middle English, and if there was a sudden upsurge in any particular period it was in the first half of the 14th century.[7]

Dictionaries record first occurrences of new words in texts regardless of their subsequent use, whereas corpora show their actual text frequencies over time. Our study did not focus on lexis but it contains two productive suffixes, *-ness* and *-ity*, whose rates of change peaked at different times in our data. While the suffixes could be seen as competing alternatives for forming abstract nouns, the normalized frequency of the native suffix *-ness* decreased sharply much earlier (between ME1–ME2) than the frequency of the borrowed *-ity* began its rise in earnest (between ME2–ME3; see Figures 15 and 16). The decline is especially prominent in deverbal *-ness*, which lacked an *-ity* counterpart and which according to Dalton-

---

7 In OED2, the total record of French loans words peaks slightly later, in the period from 1350 to 1400 (OED Online: http://www.oed.com/timelines).
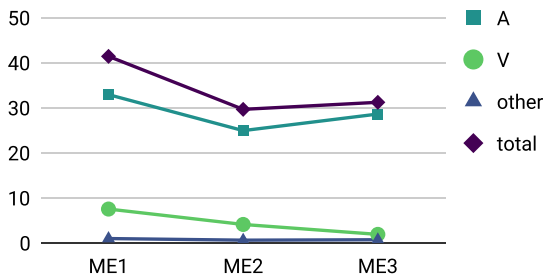
NESS

Figure 15: The (temporary) decline of -*ness* by base type (A = adjectival, V = verbal). Token frequencies from Dalton-Puffer (1992: Table 10), normalized to 10,000 words.
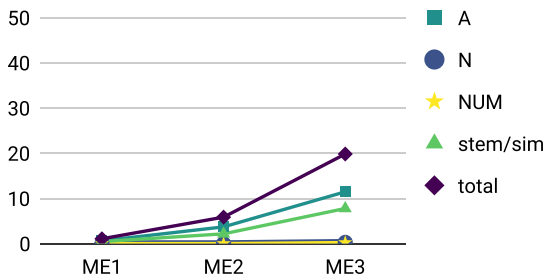
ITY

Figure 16: The rise of -*ity* by base type (A = adjectival, N = nominal, NUM = numeral, stem/sim = stem/simplex, no independent base in the data). Token frequencies from Dalton-Puffer (1996: Table 6.24), normalized to 10,000 words.

Puffer (1992: 476) had probably already started to decrease in Old English. Dalton-Puffer (1996: 84–85) hypothesizes that the decay of deverbal -*ness* could be due to its complex semantics but also to a change in the behaviour of OE versus ME translators: whereas the former tended to form morpheme-by-morpheme loan translations from Latin, the latter opted to borrow foreign words together with the concept they denoted. This could perhaps be connected to the widespread multilingualism resulting from the Conquest.

Deadjectival -*ness*, too, declined between ME1–ME2 but started to slightly increase again between ME2–ME3, which is when its competitor, deadjectival -*ity*, experienced its greatest increase. Despite the increase in -*ity*, -*ness* remained the majority option in terms of both type and token frequency throughout the ME period (Dalton-Puffer 1996: 82, 107).[8] The ME2 dip in deadjectival -*ness* could be due to multiple reasons: some of it could be explained through increased

---

8 There is also some evidence that -*ness* was increasingly used in hybrid formations combining a Romance base with the native suffix, partly inspired by French nouns ending in -*esse* (e.g. *largesse* → *largeness*), although the number of these formations remained low (Dalton-Puffer 1996: 83, 213).

competition with *-ity* and other borrowings, while some might be due to the changing composition of the corpus. The ME2 increase in the proportion of narrative texts (Figure 6), particularly romances translated or paraphrased from Anglo-Norman such as *Beues of Hamtoun*, could have led to a decrease in abstract nouns in general, and where abstract nouns were used, translators could have opted for borrowings rather than native formations as discussed above. All of this could be said to be an indirect consequence of the Conquest. In ME3, the balance of text categories is again different, and the return of secular instruction as well as the introduction of legal texts could have called for an increased use of abstract nouns.

The rate of increase of the suffix *-ity* peaks slightly later than the influx of French loanwords observed by Durkin (2014); according to Dalton-Puffer (1996: 106), it is typical of the Romance suffixes to take off in the ME3 period starting in 1350 in the *Helsinki Corpus*. In addition to the differences between lexicographical and corpus data, one explanatory factor could be that borrowing suffixes takes more time than borrowing words, as there needs to be a sufficient number of analysable words containing the suffix for it to be recognized as a suffix. However, we would argue that changes in text category balance, and indeed in text production as a whole during the period, may also have played a role in the delayed peak of the abstract noun suffix *-ity*. These Post-Invasion Effects will be discussed further in the next section.

7.2.2 Post-Invasion Effects

The Norman Conquest had far-reaching consequences on English vernacular writing. One of them was the consolidation of London as the administrative capital of the kingdom, which in the course of time impacted the dialect base of the vernacular in the written medium. Moreover, English vernacular writing was radically reduced after the Conquest, which is evidenced in the small size of the first two Middle English sections of the *Helsinki Corpus*, ME1 and ME2, and the proportion of non-contemporaneous texts in both of them (Nevanlinna et al. 1993). In this light, it is interesting to find that the major acceleration in the rate at which the recessive linguistic features in our data declined is found in the transition from ME1 to ME2. Comparing the dialect compositions of these periods shows a major shift from West Midland to East Midland dominance in ME2. It could therefore be argued that we are here witnessing a Post-Invasion Effect reflected in the dialect redistribution of the principal vernacular texts that have come down to us and excerpted in the HC.[9]

---

9 The compilers of the ME section write: "The main dialect areas are not evenly represented, but the focus changes from West to East Midlands according to the importance of the area in the extant literature of the period and the development of the language in general" (Nevanlinna et al. 1993: § 2.7).

The studies included in our material offer dialect information only on a few linguistic elements, presumably because most of the features and constructions studied are not dialect specific. There are, however, items such as the intensifier *swīþe*, which was a West-Midland form to begin with, and in fact shows its most rapid decline very early, between OE4 and ME1. This case provides support for our argument about a dialect shift following the Norman invasion in vernacular writing. It is therefore noteworthy that none of the recessive non-localized features or grammatical functions whose downward trend accelerated in ME1–ME2 went out of use but all continued into Late Middle English or, as the suffix *-ness*, into the present day. What seems to happen to *-ness*, for example, is a momentary frequency adjustment which may be in response to multiple factors ranging from other competing suffix formations to the types of text included in the corpus (see above).

Moving on to incoming features, in most cases their greatest acceleration is found at ME2–ME3. This coincides with the distinct upward turn that vernacular writing took in the second half of the 14th century in terms of both text production and text category variation. Moreover, unlike in ME2, nearly all texts in ME3 are contemporaneous. Although the distribution of regional dialects in ME2 and ME3 is more similar than that of ME1 and ME2, there are some notable differences in the text categories and proportions of verse and prose between ME2 and ME3 (see Figures 6 and 7). It could hence be argued that the post-invasion conservatism and recession of vernacular writing masked the actual rates of incoming linguistic features and that the observed acceleration in the pace of change at ME2–ME3 largely corresponds to the gradual catching up of the textual record with actual linguistic usage.

However, a further consideration to take into account is the linguistic category undergoing change. We could argue that the post-invasion delay in the documentation of processes of change is most clearly discernible in rapidly renewing intensifiers such as *most*, which soars at ME2–ME3. By contrast, the change is more evenly paced in long-term verbal processes, such as those involving verbal gerunds and the modal auxiliaries *can/could*, which only peak in ME3–ME4. These results support the view that, compared to lexical innovations, syntactic constructions are slower to change. More systematic data of both kinds would be needed to uphold this generalization but overall it is in agreement with Lass (1997: 304), who notes that stasis and punctuation are often asynchronous on different levels of language.

Finally, the text frequencies of some of the incoming features and constructions are quite low in ME3. This may be either because of their low-frequency status in general or because they are only at their incipient stage and would move on to the next, new and vigorous stage in ME4 or later (for stages of change, see

Nevalainen and Raumolin-Brunberg 2017: 54–55). If they are low-frequency items, they do not provide as robust evidence as some more frequent features. On the other hand, if a change is incipient and follows the S-curve of diffusion, it is only expected to show its greatest acceleration at a later stage. We tested this second point by comparing rates of change within a longer time span that also included the three Early Modern English periods in the HC. This test confirmed the finding that the most rapid phase of change is documented at ME2–ME3, the transition between Early and Late Middle English.

7.2.3 The role of the Black Death?

Major linguistic consequences have been attributed to the 14th-century Black Death in European languages such as Norwegian (for a critical assessment, see Mæhlum 2000). However, there has been much less discussion of the direct or indirect effects of the plague on the English language. It has been suggested, though, that the radical reduction in the population caused by the plague and the following shortage of labour promoted the prestige of the vernacular. In our data, the most rapid change at ME2–ME3 coincides with periods before and after major outbursts of the pandemic. However, while the rise of vernacular writing is strongly in evidence in the latter half of the 14th century, the possible effect of population size alone on the pace of linguistic change is harder to establish (see Section 3.3).

Historians agree that, besides the clergy, the plague hit the lower social ranks, the illiterate majority of the population, particularly hard, and that the upper, literate sections of society suffered less in terms of loss of life. Paradoxically, the situation led to a rise in popular living standards and mobility. Slack (2012: 38) summarizes the post-plague effects as follows: "the usual ties that bound farmers to their land, tenants to lords, and artisans and apprentices to their masters, all became weaker when rents and the price of land fell, wages rose, and peasants, labourers, and even servants (female as well as male) had greater bargaining power and became more mobile". Hence mobility, both social and regional, may have accelerated the pace of language change as it promoted weak ties and more loose-knit social networks (Milroy and Milroy 1985; Raumolin-Brunberg 1998). These social circumstances may be reflected in our ME3 data, especially as some of it was produced by new groups of people, such as professional secular scribes. The kinds of vernacular writing available show diachronic continuity but also renewal as secular instruction (handbooks), for example, became again prominently available in English. It is noteworthy that, while the authors of only two texts are known by name in ME2, nearly half of those in ME3 are, and they include high-

ranking individuals and professionals alike. We thus have more direct evidence for social variation in ME3 than in ME2.

In the light of this we could modify our earlier argument that ME3 represented a gradual catching up of actual linguistic change in writing to saying that ME3 was also a period of accelerated pace of linguistic change in its own right. If this is indeed the case, this acceleration cannot be presented as a direct consequence of the Black Death but rather an outcome brought about by the varied social circumstances following the pandemic. This is a provisional conclusion subject to more empirical work, and it does not preclude the caveats discussed in Section 7.2. We are suggesting it in the spirit of Slack (2012: 41–42), who acknowledges that few, if any, long-term and large-scale economic and social changes can be attributed to plague, but that the plague was nevertheless "a prime mover of undeniable force", also capable of accelerating the pace of linguistic change at least in the short term.

7.3 Assessing the results

One of the benefits of working with data from a single corpus is that the results are comparable within a given period: the data studied consist of the same set of texts. We have extended this way of thinking to major language-external forces and argued that the peaks in the change rates of features that coincide in a given period in the corpus are likely to have corresponding external explanations. While things are far from straightforward in reality, it has been possible to arrive at some preliminary conclusions about such external influences.

Overall, we found one major linguistic effect, at ME2–ME3, which marked the biggest boost in the rates of ongoing linguistic changes. However, limiting the analysis only to incoming and recessive processes yielded two such peaks, the incoming features peaking like most other changes at ME2–ME3 but the outgoing ones bunching earlier, at ME1–ME2. Both of them can be associated with external processes that were set in motion by the Norman Conquest. Hence the peaks do not manifest a direct impact of a single disruptive force, but can rather be associated with Post-Invasion Effects, an accumulation in time and space of social and cultural changes, which were mediated in writing and unfolded over a longer period of time.

Our linguistic results might have been more diffuse had the corpus data been divided into shorter time spans, but this would not have provided a solution to the relative shortage of contemporaneous material from the earliest periods, ME1 and ME2. A similar issue arises with dating the political, social and cultural forces that had an impact on linguistic processes in Middle English. For example, although

the exact dates are known of the major outbreaks of the Black Death in 1348–1349, there is no agreement among historians on the size of the English population at the time, let alone details of the movements of the population during the latter half of the 14th century (Hatcher 1994).

We could also have adduced more external evidence to support our interpretation of the data. The effects of the Black Death on the language community were probably compounded by the increased mobility of the uppermost social ranks during the Hundred Years War (see Trudgill 2020, this issue). This war was of course one of the many repercussions of the Norman Conquest – separated from the original event by some three hundred years in time.

These considerations bring us to our final point about punctuating events and Dixon's model. Many historical linguists working on reconstruction have expressed skepticism on the sharp dichotomy into equilibrium and punctuating events suggested by the model. Campbell (2003: 50–51), for one, points out that, unlike what the model predicts, linguistic divergence also takes place in equilibrium, and so does convergence in punctuation. He refers to the history of English as mostly one of punctuation, shaped by the Scandinavian invasion and the Norman Conquest, and suggests that the outcome in fact resembles more what one would expect from equilibrium states: massive vocabulary assimilation, borrowed sounds and pronouns, and morphosyntactic levelling.

Such high-level generalizations are of course partly a matter of granularity of the time scale adopted. Our results support the view that punctuating events such as the Norman Conquest could have specific long-term consequences arising from diverse local and contextual effects on the language community. Although these effects could impact the rate of linguistic changes, they rarely interrupted them altogether. If these Post-Invasion Effects are thought to have materialized in a period of (relative) stability, our evidence would support the original notion of linguistic convergence in equilibrium.

8 Conclusion

With this study we have taken the first steps towards answering some fundamental questions about the nature of language change by carrying out a meta-analysis of a number of linguistic changes in the history of English. Our study was facilitated by a new research database, the LCD, which provided us with the numerical data required for such an undertaking, and our decision to focus on research done with the *Helsinki Corpus* ensured that the data drawn from different studies remained comparable. The research design allowed us to put existing hypotheses concerning the rate of language change to the test and reach some preliminary conclusions

about the way major historical punctuations accelerate change. We would like to emphasize, however, that while our data show clear trends both within the Middle English period and the longer diachrony, the results should be re-examined in future research against a more ample body of data: as one of the compilation principles of the LCD was to make earlier corpus-linguistic research better available to the research community (Nevalainen et al. 2016), much of the data used in the present study was drawn from research done in the 1990s, and the most recent research is therefore somewhat underrepresented in our analysis.

The method we have developed for analysing the rate of change is likewise a preliminary one. Here our main criteria were those of simplicity and transparency: rather than crafting a complex statistical model whose background assumptions might not be met by the data and whose operation might be somewhat opaque, we wanted to devise a method that would be robust and easy to understand, with results that would be easy to interpret. While in this paper we focused on normalized text frequencies, the method could also be applied to the proportion of a variant out of a linguistic variable, which might be a more suitable alternative for some changes. The robustness of the method could be further augmented by implementing a bootstrapping technique to calculate confidence intervals. Furthermore, modelling frequency change beyond the observed data, as in Van de Velde (2017: 70–72), could provide us with an additional tool for estimating the period of fastest change across multiple changes.

One theoretical question that our study raises is when a given period of punctuation ends and one of equilibrium begins. Our analysis was based on data that covers the period of Middle English from 1150 to 1500, and it showed one particular juncture at which remarkably many linguistic processes displayed the most change, i.e. between ME2 and ME3, some three hundred years after the Norman Conquest in 1066; in the history of English, this juncture also coincides with the transition from Early to Late Middle English. Our overall results thus suggest that the effects of the Conquest on the pace of linguistic change were indirect and deferred. In ME3, the period 1350–1420, this outcome must also have been partly influenced by the other major punctuating event of the Black Death. These observations lead us to conclude that, as far as the Norman Conquest is concerned, the acceleration of changes that we have observed comes in response to its after-effects in a period of (relative) stability. However, the overall outcome of our exploration supports previous research: the documented history of Middle English is marked by accelerated linguistic change.

References

Primary sources

Dalton-Puffer. Christiane. 1992. The status of word formation in Middle English: Approaching the question. In Matti Rissanen, Ossi Ihalainen, Terttu Nevalainen & Irma Taavitsainen (eds.). *History of Englishes: New methods and interpretations in historical linguistics (Topics in English Linguistics 10)*, 465–482. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110877007.465.

Dalton-Puffer, Christiane. 1996. Abstract noun suffixes. In *The French influence on Middle English morphology: A corpus-based study of derivation* (Topics in English Linguistics 20), 73–130. Berlin: Mouton de Gruyter.

De Smet, Hendrik. 2008. Functional motivations in the development of nominal and verbal gerunds in Middle and Early Modern English. *English Language and Linguistics* 12(1). 55–102.

Iyeiri, Yoko. 2002. Development of *any* from Middle English to Early Modern English: A study using the Helsinki Corpus of English Texts. In Toshio Saito, Junsaku Nakamura & Shunji Yamazaki (eds.), *English corpus linguistics in Japan (Language and Computers: Studies in Practical Linguistics 38)*, 211–223. Amsterdam: Rodopi. https://doi.org/10.1163/9789004334205_013.

Kahlas-Tarkka, Leena. 1993. Toward the Modern English dichotomy between *every* and *each*. In Matti Rissanen, Merja Kytö & Minna Palander-Collin (eds.), *Early English in the computer age: Explorations through the Helsinki Corpus (Topics in English Linguistics 11)*, 201–218. Berlin: Mouton de Gruyter.

Kilpiö, Matti. 1997. Participial adjectives with anaphoric reference of the type *the said, the (a)forementioned* from Old to Early Modern English: The evidence of the Helsinki Corpus. In Terttu Nevalainen & Leena Kahlas-Tarkka (eds.), *To explain the present:*

Studies in the changing English language in honour of Matti Rissanen (Mémoires de la Société Néophilologique de Helsinki 52), 77–101. Helsinki: Société Néophilologique.

Kytö, Merja. 1991. *Can (could)* vs. *may (might)* in Old and Middle English: Testing a diachronic corpus. In *Variation and diachrony, with Early American English in focus (Bamberger Beiträge zur Englischen Sprachwissenschaft 28)*, 131–180. Frankfurt am Main: Peter Lang.

Loureiro-Porto, Lucía. 2008. The convergence of two *need* verbs in Middle English. In Richard Dury, Maurizio Gotti & Marina Dossena (eds.), *English Historical Linguistics 2006: Selected papers from the Fourteenth International Conference on English Historical Linguistics (ICEHL 14)*, vol. II: *Lexical and semantic change (Current Issues in Linguistic Theory 296)*, 97–116. Amsterdam: John Benjamins.

Méndez-Naya, Belén. 2003. On intensifiers and grammaticalization: The case of *swīþe*. *English Studies* 84(4). 372–391.

Méndez-Naya, Belén. 2006. Adjunct, modifier, discourse marker: On the various functions of *right* in the history of English. *Folia Linguistica Historica* 27(1–2). 141–195.

Méndez-Naya, Belén. 2008. *The which is most and right harde to answere*: Intensifying *right* and *most* in earlier English. In Richard Dury, Maurizio Gotti & Marina Dossena (eds.), *English Historical Linguistics 2006: Selected papers from the Fourteenth International Conference on English Historical Linguistics (ICEHL 14)*, vol. II: *Lexical and semantic change (Current Issues in Linguistic Theory 296)*, 31–51. Amsterdam: John Benjamins.

Österman, Aune. 1997. *There* compounds in the history of English. In Matti Rissanen, Merja Kytö & Kirsi Heikkonen (eds.), *Grammaticalization at work: Studies in long-term developments in English (Topics in English Linguistics 24)*, 191–276. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110810745.191.

Österman, Aune. 2001. From *there* to *where*: The development of relative and conjunctive adverbs in Middle English. *NOWELE* 38(1). 65–107.

Palander-Collin, Minna. 1997. A medieval case of grammaticalization, *methinks*. In Matti Rissanen, Merja Kytö & Kirsi Heikkonen (eds.), *Grammaticalization at work: Studies of long-term developments in English (Topics in English Linguistics 24)*, 371–404. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110810745.371.

Peitsara, Kirsti. 1997a. *Enough* and *enow* in Middle English. In Terttu Nevalainen & Leena Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen (Mémoires de la Société Néophilologique de Helsinki 52)*, 163–183. Helsinki: Société Néophilologique.

Peitsara, Kirsti. 1997b. The development of reflexive strategies in English. In Matti Rissanen, Merja Kytö & Kirsi Heikkonen (eds.), *Grammaticalization at work: Studies of long-term developments in English (Topics in English Linguistics 24)*, 277–370. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110810745.277.

Raumolin-Brunberg, Helena & Leena Kahlas-Tarkka. 1997. Indefinite pronouns with singular human reference. In Matti Rissanen, Merja Kytö & Kirsi Heikkonen (eds.), *Grammaticalization at work: Studies of long-term developments in English (Topics in English Linguistics 24)*, 17–86. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110810745.17.

Rissanen, Matti. 1997. Optional THAT with subordinators in Middle English. In Raymond Hickey & Stanisław Puppel (eds.), *Language history and linguistic modelling: A Festschrift for Jacek Fisiak on his 60th birthday (Trends in Linguistics: Studies and Monographs 101)*, 373–383. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110820751.373.

Secondary sources

Aitchison, Jean. 1981. *Language change: Progress or decay?* London: Fontana Paperbacks.

Altmann, Eduardo G, Janet B. Pierrehumbert & Adilson E. Motter. 2009. Beyond word frequency: Bursts, lulls, and scaling in the temporal distributions of words. *PLoS ONE* 4(11). e7678.

Bailey, Charles-James N. & Karl Maroldt. 1977. The French lineage of English. In Jürgen M. Meisel (ed.), *Langues en contact: Pidgins, creoles/Languages in contact*, 21–53. Tübingen: Narr.

Bartlett, Robert. 2000. *England under the Norman and Angevin Kings*, 1075–1225 (New Oxford History of England, gen. ed. John M. Roberts). Oxford: Clarendon Press.

Bergs, Alexander. 2005. *Social networks and historical sociolinguistics: Studies in morphosyntactic variation in the Paston Letters (1421–1503)*. Berlin & New York: Mouton de Gruyter.

Bergs, Alexander & Laurel J. Brinton (eds.). 2012. *Historical linguistics of English (HSK 34.1–2)*. Berlin & New York: de Gruyter Mouton.

Biber, Douglas & Susan Conrad. 2009. *Register, genre and style*. Cambridge: Cambridge University Press.

Britnell, Richard. 2006. Town life. In Rosemary Horrox & W. Mark Ormrod (eds.), *A social history of England, 1200–1500*, 134–178. Cambridge: Cambridge University Press.

Campbell, Lyle. 2003. Beyond the comparative method? In Barry J. Blake & Kate Burridge (eds.), assisted by Jo Taylor, *Historical linguistics 2001: Selected papers from the 15th International Conference on Historical Linguistics, Melbourne, 13–17 August 2001*, 33–58. Amsterdam & Philadelphia: John Benjamins.

Cooper, Harris, Larry V. Hedges & Jeffrey C. Valentine (eds.). 2009. *The handbook of research synthesis and meta-analysis*, 2nd edn. New York: Russell Sage Publication.

Crystal, David. 2003. *The Cambridge encyclopedia of the English language*, 2nd edn. (3rd edn. 2019). Cambridge: Cambridge University Press.

Dalton-Puffer, Christiane. 1996. *The French influence on Middle English morphology: A corpus-based study of derivation (Topics in English Linguistics 20)*. Berlin: Mouton de Gruyter.

Denison, David. 2003. Log(ist)ic and simplistic S-curves. In Raymond Hickey (ed.), *Motives for language change*, 54–70. Cambridge: Cambridge University Press.

Dixon, Robert M. W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.

Durkin, Philip. 2014. *Borrowed words: A history of loanwords in English*. Oxford: Oxford University Press.

Durrant, Philip. 2014. Corpus frequency and second language learners' knowledge of collocations. *International Journal of Corpus Linguistics* 19(4). 443–477.

Eldredge, Niles & Stephen J. Gould. 1972. Punctuated equilibria: An alternative to phyletic gradualism. In Thomas J. M. Schopf (ed.), *Models in paleobiology*, 82–115. San Francisco: Freeman Cooper.

Emonds, Joseph E. & Jan Terje Faarlund. 2014. *English: The language of the Vikings (Olomouc Modern Language Monographs 3.)* Olomouc: Palacký University.

Filppula, Markku & Juhani Klemola (eds.). 2009. Re-evaluating the Celtic Hypothesis. Special issue of *English Language and Linguistics* 13(2).

Goo, Jaemyung, Gisela Granena, Yucel Yilmaz & Miguel Novella. 2015. Implicit and explicit instruction in L2 learning: Norris & Ortega (2000) revisited and updated. In Patrick Rebuschat (ed.), *Implicit and explicit learning of languages*, 443–482. Amsterdam & Philadelphia: John Benjamins.

Greenhill, Simon J, Xia Hua, Caela F. Welsh, Hilde Schneemann & Lindell Bromham. 2018. Population size and the rate of language evolution: A test across Indo-European, Austronesian, and Bantu languages. *Frontiers in Psychology* 9. 576.

Hatcher, John. 1994. *Plague, population and the English economy 1348–1530*. Houndmills & London: Macmillan Press.

HC = *The Helsinki Corpus of English Texts*. 1991. Compiled by Matti Rissanen (Project leader), Merja Kytö (Project secretary); Leena Kahlas-Tarkka, Matti Kilpiö (Old English); Saara Nevanlinna, Irma Taavitsainen (Middle English); Terttu Nevalainen, Helena Raumolin-Brunberg (Early Modern English). Department of Modern Languages, University of Helsinki.

Higham, Nicholas & Martin J. Ryan. 2013. *The Anglo-Saxon world*. New Haven & London: Yale University Press.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Ingham, Richard. 2014. The maintenance of French in later medieval England. *Neuphilologische Mitteilungen* 115(4). 623–645.

Ingham, Richard. 2017. English and French in medieval England: Spoken bilingualism or code diglossia? In Elise Louviot & Catherine Delesse (eds.), *Studies in language variation and change 2: Shifts and turns in the history of English*, 175–196. Newcastle upon Tyne: Cambridge Scholars Publishing.

Kahlas-Tarkka, Leena, Matti Kilpiö & Aune Österman. 1993. Old English. In Matti Rissanen, Merja Kytö & Minna Palander (eds.), *Early English in the computer age: Explorations through the Helsinki Corpus*, 21–32. Berlin & New York: Mouton de Gruyter. http://www.helsinki.fi/varieng/CoRD/corpora/HelsinkiCorpus/oldenglish.html (accessed 18 December 2019).

Kauhanen, Henri & George Walkden. 2018. Deriving the constant rate effect. *Natural Language & Linguistic Theory* 36(2). 483–521.

Keen, Maurice. 1990. *English society in the later Middle Ages 1348–1500*. London: Penguin Books.

Keene, Derek. 2000. Metropolitan values: Migration, mobility and cultural norms, London 1100–1700. In Laura Wright (ed.), *The development of Standard English 1300–1800: Theories, descriptions, conflicts*, 93–114. Cambridge: Cambridge University Press.

Kesäniemi, Joonas, Turo Vartiainen, Tanja Säily & Terttu Nevalainen. 2018. Exploring meta-analysis for historical corpus linguistics based on linked data. *Journal of Research Design and Statistics in Linguistics and Communication Science* 5(1–2). 4–47.

Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe – Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36. 15–43.

Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1(3). 199–244.

Kytö, Merja & Matti Rissanen. 1983. The syntactic study of early American English: The variationist at the mercy of his corpus? *Neuphilologische Mitteilungen* 84(4). 470–490.

Lass, Roger. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.

Lijffijt, Jefrey, Terttu Nevalainen, Tanja Säily, Panagiotis Papapetrou, Kai Puolamäki & Heikki Mannila. 2016. Significance testing of word frequencies in corpora. *Digital Scholarship in the Humanities* 31(2). 374–397.

Mæhlum, Brit. 2000. Social catastrophes as explanation in historical linguistics. In Ernst Håkon Jahr (ed.), *Språkkontakt: Innverknaden frå nedertysk på andre nordeuropeiske språk*, 87–94. Copenhagen: Nordisk Ministerråd.

Miller, Gary D. 2012. *External influences on English: From its beginnings to the Renaissance*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199654260.001.0001.

Milroy, James & Lesley Milroy. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21. 339–384.

Mugglestone, Lynda (ed.). 2012. *The Oxford history of English*, rev. edn. Oxford: Oxford University Press.

Nevalainen, Terttu. 2015. Descriptive adequacy of the S-curve model in diachronic studies of language change. In Christina Sanchez-Stockhammer (ed.), *Can we predict linguistic change? (Studies in Variation, Contacts and Change in English 16)*. Helsinki: VARIENG. http://www.helsinki.fi/varieng/series/volumes/16/nevalainen/ (accessed 18 December 2019).

Nevalainen, Terttu & Elizabeth Closs Traugott (eds.). 2012. *The Oxford handbook of the history of English*. New York & Oxford: Oxford University Press.

Nevalainen, Terttu & Helena Raumolin-Brunberg. 2017. *Historical sociolinguistics: Language change in Tudor and Stuart England*, 2nd edn. London & New York: Routledge.

Nevalainen, Terttu, Turo Vartiainen, Tanja Säily, Joonas Kesäniemi, Agata Dominowska & Emily Öhman. 2016. Language Change Database: A new online resource. *ICAME Journal* 40. 77–94.

Nevanlinna, Saara, Päivi Pahta, Kirsti Peitsara & Irma Taavitsainen. 1993. Middle English. In Matti Rissanen, Merja Kytö & Minna Palander (eds.), *Early English in the computer age: Explorations through the Helsinki Corpus*, 33–51. Berlin & New York: Mouton de Gruyter. http://www.helsinki.fi/varieng/CoRD/corpora/HelsinkiCorpus/middleenglish.html (accessed 18 December 2019).

Norris, John M. & Lourdes Ortega (eds.). 2006. *Synthesizing research on language learning and teaching*. Amsterdam: John Benjamins.

OED = *Oxford English Dictionary*. OED Online. Oxford: Oxford University Press. https://www.oed.com (accessed 18 December 2019).

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Raumolin-Brunberg, Helena. 1998. Social factors and pronominal change in the seventeenth century: The Civil-War effect? In Jacek Fisiak & Marcin Krygier (eds.), *Advances in English historical linguistics (1996)*, 361–388. Berlin: Mouton de Gruyter.

Renaud, Jean. 2008. The Duchy of Normandy. In Stefan Brink & Neil Price (eds.), *The Viking world*, 453–457. Abingdon: Routledge.

Samuels, Michael L. 1972. *Linguistic evolution: With special reference to English*. Cambridge: Cambridge University Press.

Schendl, Herbert. 2012. Middle English: Language contact. In Alexander Bergs & Laurel J. Brinton (eds.), *Historical linguistics of English (HSK 34.1)*, 505–519. Berlin & New York: De Gruyter Mouton.

Sinnemäki, Kaius. 2020. Linguistic system and sociolinguistic environment as competing factors in linguistic variation: A typological approach. *Journal of Historical Sociolinguistics*. In this issue.

Slack, Paul. 2012. *Plague: A very short introduction*. Oxford: Oxford University Press.

Strohm, Paul. 2006. Writing and reading. In Rosemary Horrox & W. Mark Ormrod (eds.), *A social history of England, 1200–1500*, 454–472. Cambridge: Cambridge University Press.

Thomas, Hugh M. 2008. *The Norman Conquest: England after William the Conqueror*. Lanham, Boulder & New York: Rowman & Littlefield Publishers.

Thomason, Sarah G. 2008. Social and linguistic factors as predictors of contact-induced change. *Journal of Language Contact*, THEMA series, Number 2, Language Contact and the Dynamics of Language. 42–56.

Timofeeva, Olga & Richard Ingham. 2018. Introduction. Special issue on mechanisms of French contact influence in Middle English: diffusion and maintenance. In Olga Timofeeva & Richard Ingham (eds.), *English Language and Linguistics* 22(2). 197–205.

Townend, Matthew. 2012. Contacts and conflicts: Latin, Norse, and French. In Lynda Mugglestone (ed.), *The Oxford history of English, rev. edn*, 75–105. Oxford: Oxford University Press.

Trudgill, Peter. 2016. Contact-related processes of change in the early history of English. In Merja Kytö & Päivi Pahta (eds.), *The Cambridge handbook of English historical linguistics*, 318–334. Cambridge: Cambridge University Press.

Trudgill, Peter. 2020. Sociolinguistic typology and the speed of linguistic change. *Journal of Historical Sociolinguistics*. In this issue.

Van de Velde, Freek. 2017. Understanding grammar at the community level requires a diachronic perspective: Evidence from four case studies. *Nederlandse Taalkunde* 22(1). 47–74.

Warner, Anthony. 2017. English–Norse contact, simplification and sociolinguistic typology. *Neuphilologische Mitteilungen* 118(2). 317–403.