Sociocultural Dimensions of Lexis and Text in the History of English

EDITED BY

Peter Petré

Hubert Cuyckens

Frauke D'hoedt

JOHN BENJAMINS PUBLISHING COMPANY

SOCIOCULTURAL DIMENSIONS OF LEXIS AND TEXT
IN THE HISTORY OF ENGLISH

CURRENT ISSUES IN LINGUISTIC THEORY

General Editor

JOSEPH C. SALMONS

University of Wisconsin–Madison

jsalmons@wisc.edu

Founder & General Editor (1975-2015)

E.F.K. KOERNER

Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin

efk.koerner@rz.hu-berlin.de

Current Issues in Linguistic Theory (CILT) is a theory-oriented series which welcomes contributions from scholars who have significant proposals that advance our understanding of language, its structure, its function and especially its historical development. CILT offers an outlet for meaningful contributions to current linguistic debate.

A complete list of titles in this series can be found on *http://benjamins.com/catalog/cilt*

Volume 343

Peter Petré, Hubert Cuyckens and Frauke D'hoedt (eds.)

Sociocultural Dimensions of Lexis and Text in the History of English

SOCIOCULTURAL DIMENSIONS OF LEXIS AND TEXT IN THE HISTORY OF ENGLISH

Edited by

PETER PETRÉ

University of Antwerp

HUBERT CUYCKENS

KU Leuven

FRAUKE D'HOEDT

KU Leuven

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM & PHILADELPHIA

Table of contents

Foreword

The chapters in the present volume spring forth from work that was first presented at the 18th *International Conference of English Historical Linguistics* (ICEHL18), held at the University of Leuven, 14–18 July, 2014. ICEHL, which is held in Europe biennially, has alternated since 1992 with the International Conference of Historical Linguistics. Its long tradition goes back to 1979, when it was organized for the first time by the University of Durham. Yet it had never been organized in Belgium before. It is no coincidence that the University of Leuven should be the first Belgian university to have done so. After all, English historical linguistics has been at the heart of its linguistics department ever since Xavier Dekeyser started a History of English class in 1971.

For turning the conference into a success, we are first of all obliged to all the presenters and conference participants, who created a stimulating atmosphere of intellectual exchange. Special thanks also go to the plenary speakers Charles Boberg, Robert Fulk, Peter Grund, María José López-Couso, and Marit Westergaard, who demonstrated that English historical linguistics is truly a multifaceted research domain. We would furthermore like to thank our co-organizers Hendrik De Smet and Liesbet Heyvaert, Lauren Fonteyn, Charlotte Maekelberghe, and Nikki van de Pol, and all our student helpers, as well as Lieselotte Brems (Université de Liège). For financial support, we are grateful to the Research Foundation Flanders (FWO).

There is a longstanding tradition of publishing a selection of ICEHL papers with John Benjamins, and we are happy to continue this tradition. To ensure thematic coherence, accepted submissions were divided into two volumes. The present one, which is published in the time-honored series *Current Issues in Linguistic Theory*, focuses on the cultural embedding of change. A companion volume entitled *Explorations in English historical syntax*, is to appear in another venerable series by Benjamins, *Studies in Language Companion Series*. For the present volume we have made a careful selection from a generous number of submissions. This would not have been possible without all of the detailed review reports, for which we would like to thank all the reviewers: Birgit Alber, Kathryn Allan, Renate Bauer, Cristiano Broccias, Luisella Caon, Claudia Claridge, Jonathan Culpeper, Hans-Jürgen Diller, Marina Dossena, Robert Fulk, Peter Grund, Ryuichi Hotta, Joan Houston Hall, Luis Iglesias-Rabade, Elisa Mattiello, Lynda Mugglestone, Leonard Neidorf, Elisabeth Okasha, Päivi Pahta, David Parsons, Carol Percy, Javier Pérez Guerra, Betty

Phillips, Nikolaus Ritt, Sharon Rowley, Ana Sáez-Hidalgo, Roland Schuhmann, Jeremy Smith, Louise Sylvester, Harumi Tanabe, David Tizón Couto, Ferdinand von Mengden, Brita Warvik, and Jerzy Wełna. Many thanks also go to Anke de Looper of John Benjamins for seeing this project through and for her friendly advice. Finally we would also like to express our thanks to the series editor, Joe Salmons, and his editorial team, for accepting the manuscript for publication in their series, and for their careful screening of the manuscript and helpful comments.

<div align="right">Peter Petré, Hubert Cuyckens & Frauke D'hoedt</div>

Introduction

Philology as linguistically informed cultural history

Peter Petré and Hubert Cuyckens

University of Antwerp / KU Leuven

The present volume collects ten studies that can be situated on the philological crossroads between literature and linguistics. Philological research has always been an integral part of Germanic language studies. Among other areas, English historical philology includes such time-honored subfields as dialectology and paleography (for an overview see, for instance, Laing et al. 1989; Hogg 2009; Laing & Lass 2009) and standardization (see, for instance, Stein & Tieken-Boon van Ostade 1993). It has also seen, since the 1980s, the emergence of newer subfields such as historical sociolinguistics, often with special attention to social networks (see, for instance, Nevalainen & Raumolin-Brunberg 2017).

The viewpoint complementary to these themes, which is shared by all contributions in this volume, is a *sociohistorical* one. Specifically, the studies link various linguistic dimensions of lexis and morphology to the sociohistorical and cultural contexts of the time (especially those that are not at the core of social network research). Each paper pays particular attention to social 'labels' and attitudes (conservative, religious, ideological, endearing, or other), thereby focusing on the dynamic and historical dimension of these attitudes and labels. Changes in these are analyzed in order to explain linguistic changes that would otherwise be hard to account for. The wide range of linguistic phenomena covered demonstrates how this approach can shed new light on language change at all levels. The effect of changing attitudes on morphology is illustrated with the changes in word-clipping tendencies discussed by Minkova in Chapter 10. At the lexical level, attitudinal changes are exposed in the discussion by Boiché of Anglo-Saxon naming practices in Chapter 1, and the attraction of a new attitudinal layer is exemplified in Chapter 9 (by Hotta), which is on the increased association of prepositional variants in -*st* with the label 'archaic'. Suhr's examination of the visual dimension of early modern news pamphlets (Chapter 2) illustrates how changing attitudes may also affect the balance of extra-linguistic and linguistic features of a text.

© 2018 John Benjamins Publishing Company

Multiple layers of different attitudes are also important to understand linguistic change in the context of text reproduction. Rather than adhering to traditional philology (including its paleographical and dialectological dimensions), whose purpose has been the identification and characterization of the 'authoritative version', the present volume subscribes to the so-called 'new philology' or 'material philology' (Nichols 1997), which has been influential in literary philology for a while, but is less commonly adopted in linguistically oriented philology. This 'material philology' has decentralized the concept of 'original authorship' and, instead, has made the manuscript itself central. The chapters that deal with text reproduction in this volume all take this approach. In Chapter 4, Wallis discusses what the behavior of a copyist tells us about that copyist. While copyists attempt to faithfully render the exemplar they are copying, a copy is rarely if ever identical to its exemplar, and therefore consists of at least two layers of textuality: the layer contributed by the author and that contributed by the scribe. Each layer is meaningful and tells us something about the attitudes of its creator. Another example is Gather's analysis of congregational songs as early modern adaptations of older psalm translations (Chapter 7). Multiple layers may even be found in different editions of the same text, as is shown in Ruano-García's analysis of the dictionary of Americanisms (Chapter 8), which reveals how an editor deals with a changing view on what constitutes American English.

In addition to the sociohistorical perspective, (English) philological research is also being greatly affected by what is commonly referred to as 'Digital Humanities', an umbrella term that covers a wide range of methodologies, ranging from straightforward computer-assisted analysis of texts to powerful computer software to analyze and visualize enormous amounts of data. While the surge of Digital Humanities and Big Data is a welcome and highly relevant boost to the field of philology, it also constitutes a danger to the richness of the philological tradition. Large-scale quantitative analyses inherently run the risk of undervaluing differences in the data, and of losing sight of the underlying message of individual texts and authors. These potential risks do not mean that we should shy away from quantitative analysis. Ideally, both quantitative and qualitative analysis will be combined in a thoughtful way. The current volume explores such thoughtful combination. While it capitalizes on the hallmarks of traditional philological craftsmanship, the majority of its contributions combine this craftsmanship with the opportunities provided by computer-assisted text analysis.

Several of the papers embark on a quantitative analysis of semantic, pragmatic, or syntactic distinctions in historical texts, against this shared background of the texts' sociohistorical markup. While quantitative analysis is becoming more and more common, quantification still remains very much a non-trivial undertaking when investigating philological topics. The attention that is given in this volume

to quantification in the discussion of these philological topics is therefore a wel-come complement to the qualitative interpretation of the data, which still forms the cornerstone of this volume. This attention includes exploratory quantitative analyses of topics that are particularly hard to quantify such as metaphor research (Alexander & Kay in Chapter 3), that simply have not yet received a quantitative treatment (Moessner on the legal genre in early English, Chapter 5), or that try to generalize over quantitative evidence from different syntactic domains (e.g. mood, tense, and person deixis; Nakayasu in Chapter 6). Other chapters are predominantly qualitative in nature, but raise important issues with regard to alternative quanti-tative approaches. For example, Wallis' analysis of scribe–exemplar relationships (Chapter 4) reveals that the dialect difference between the exemplar and the copy, with each dialect having its own set of distinctive phonological features, is a likely source of a scribe's confusion as manifested in unexpected interpretative alterations of his exemplar. Such differences are not yet fully taken into account in the recent surge of computational stylometric research trying to discriminate between copyist and author features (e.g. Kestemont & Van Dalen-Oskam 2010; Kestemont et al. 2015).

The volume is divided into three main parts, each focusing on a different per-spective of the sociocultural dimensions of lexis and text. The first perspective, entitled *Conspicuous lexical choice in past societies*, focuses on how **meaning change is affected by a changing society**. The case studies depart from well-studied top-ics such as the impact of the Danish and Anglo-Norman invasions on the Old and Middle English (ME) language or the rise of prescriptivism in the English renaissance; rather, they discuss, *inter alia*, the social mobility of proper names independent of their owners, the effects of increased literacy on news headings, and the past habit of using the supernatural as a (metaphorical or literal?) excuse to account for undesirable situations.

The second part, which has been labeled *Historical layers in text and genre*, focuses on **how genres and text types may be enriched by the attitudes of their producers**. Analyses in this part start from the well-known assumption that genres and registers can only be identified by contrasting them with other genres (Biber & Conrad 2009): they compare various legal genres, or contrast non-fiction and fiction by the same author. Not only do genres differ among each other, but one can also not simply assume that they remain stable over time. The systematic study of how genres change and are affected (for instance, by register) is an area where much work remains to be done (see Biber & Gray 2013). All papers in this part reveal changes within genres over time, both genres that no longer exist today, such as congregational song, and genres that are still in use today. Within the category of characteristics that have been lost from certain genres, we may mention such un-expected ones as curse clauses in wills, the almost exclusively negative connotation

of the concept of *news* – admittedly, sensationalist news still sells well today – or the fact that medieval handbooks were often structured as dialogues.

The third part, *Lexis and morphology in a changing society*, reveals **how affective labels** such as 'Americanism', 'archaicism' (of certain prepositions ending in *-st*), or 'sloppy language use' (of foreclippings such as *minish* for *diminish*) **may become associated with formerly neutral vocabulary**. Specifically, the contributions to this part show how such labeling is in a way a by-product of more encompassing changes, respectively the emergence of an American identity (not merely linguistic but also cultural), the loss of the *-st* morpheme elsewhere in the English language (think of second-person singular verb agreement), or the decline of a monolithic Germanic stress constraint due to massive language contact with French.

The three following sections briefly introduce the contents of each part in turn.

Conspicuous lexical choice in past societies

Part 1 contains three papers dealing with lexis. Each examines lexical choice as the coding of speakers' subjective categorization of the outside world, and in particular of individuals and groups that are somehow perceived as special, and have accordingly received various 'social labels'. Given the pre-secularized sociohistorical context of the papers, this special status persistently assumes a religious or spiritual dimension.

The first paper (Boiché) reveals the sociocultural motivations underlying Old English names, which were as a rule much more meaningful than in today's English-speaking world, where name-giving is more often motivated purely by a name's euphonic qualities. One Anglo-Saxon onomastic morpheme, *ead-*, is analyzed in particular detail, and it is also compared to another (*wela-*) and similar morphemes in Greek and Slavic. Boiché provides evidence that the earliest attestations of the lexical element *ead* (*aud* in common Germanic) were independent uses as a name meaning "prosperity", and were given to procure a bright future for the newborn. She argues that at this stage, this name was probably typical of the lower social stratum. Only later on, from the 5th century onwards, did it spread to gentry. This spread was triggered by *ead* being used as the first element in dithematic names, such as *Odovacer* "watchman of property". *Ead*, subsequently, became very productive as the first element of such names. Boiché's argument effectively suggests a fascinating relationship between the social mobility of a particular name and its changing morphological status, from being independent to becoming a bound morpheme in complex names.

The second paper (Suhr) examines the pragmatics of the textual and visual features of early modern sensationalist news titles dealing with monsters (i.e.

monstrous births), storms, and the devil. The analysis shows that some textual labels were preferred over others. Occurrences of *news*, with its clear emphasis on revealing the unknown, outnumbered *history*, a situation which persists to this day, and *relations* was favored over a label such as *account*. Yet regardless which label was used, textual labels were not generally typographically highlighted until the 1640s. Before that time the overall visual layout of the title page was a more important genre marker than textual labels. With regard to readership, sensationalist news pamphlets, unlike learned polemical pamphlets, had an increasingly wide audience consisting of growing numbers of newly semi-literate readers from all but the lowest levels of society. The developments in the conventionalization of textual labels – and its prevalence over the visual layout – can be tied in with increasing literacy and access to printed texts of these unlearned masses. Lexical and textual markup decisions in this case reflect the changing sociocultural situation of the audience.

The final paper of Part 1 (ALEXANDER & KAY) examines the sociocultural background of the metaphorical use of lexis referring to creatures from the spirit world in the past. Based on the large-scale analysis of metaphorical mappings, the authors observe that reference to the supernatural could have good (*angelic, heavenly*) or bad (*impish*) connotations, but that negative overtones are clearly predominant. They link this to the tendency of humans to appeal to supernatural interference to explain away their evils and misfortunes (think of devils, nightmares, elves, imps, etc.), but rarely to refer to their virtues and successes.[1] In addition to providing a quantitative analysis of metaphorical domains, still an underexplored methodology in metaphor research, Alexander & Kay's paper also argues that metaphorical mappings such as those from the supernatural to the human provide a conceptual challenge for metaphor research. In contemporary cognitive semantics, metaphor is generally treated as the mapping of a concrete source domain to a(n often more abstract) target domain. It could be argued though, that supernatural powers are more abstract than their 'metaphorical' human targets. But note that the abstract nature of the supernatural may be an anachronistic projection of our own worldview. Heaven, as the abode of God, for instance, may have been considered to be a physical place above our heads, and as such polysemous to its meaning of sky and equally concrete. In this view, speakers would have applied standard metaphorical strategies when calling a concrete place *heaven* (e.g. "like heaven"). In today's perspective, such a mapping would be unexpected, as it would be from more abstract to more concrete. Similarly, the concreteness of the supernatural concept of *incubus*

---

1.   This is an observation that nicely ties in with research on memory performance: people seem to remember their virtuous actions better than their immoral ones (e.g. Kouchakia & Ginob 2016).

"evil male spirit", the source of metaphors on human diseases, may be derived from its being mentioned as an entity in medieval law. Cognitive semantics should therefore take into account the importance of the historical and sociocultural context in semantic networks. At the same time, this concreteness of the supernatural can hardly be equated with the palpable concreteness of the physical world. The concept of metaphor as conceived in contemporary cognitive semantics itself may be a historical one: its definition as being a similarity between a concrete source and a more abstract target makes use of a distinction (concrete–abstract) which itself only became entrenched in society in the 20th century. The authors propose to split up the concept of metaphor into a metaphorical world accepted as literal in some belief systems, and 'pure' metaphors.

Historical layers in text and genre

The four papers in Part 2 center on the social and cultural dimensions at the level of the text and genre. In particular, they show that linguistic analysis – orthographic as well as semantic and syntactic – can provide a window on how the identity of texts and genres is tied to sociocultural attitudes towards the past.

The first paper (WALLIS) evaluates the layer of interventions added by scribe B1 of the Corpus Christi copy of *Bede* (CCCC 41), when he was faced with an exemplar that was written about 150 years before in a different dialect. One aspect that is particularly highlighted in this contribution is that the modernizations of scribes are not only interesting as evidence of diachronic change, but they may reveal much about the motivations, methods, and practices of the scribes who produced our surviving witnesses. This is in contrast to the majority of studies, which typically focus on what later transmission can reveal about lost archetypes and exemplars. Wallis shows that B1 is not the careless scribe he has sometimes been seen as (Grant 1989: 10). Instead, the nature of his interventions reflects the intelligent approach of a translator scribe who, even when faced with language too different from his own to be easily understood, pays due respect to his exemplar. Unfamiliar phrases and lexis are regularly replaced by West-Saxon alternatives. While these substitutions are not always successful, such as *Hibernia* "Ireland" for *hii* (the isle *Iona*), they clearly try to combine some of the original phonological shape with appropriate meaning. B1 also adopts some paleographical and morphological features of his exemplar, but only if they do not hamper the interpretation. Anglian double vowels are generally replaced by West-Saxon spelling conventions, except when they occur in words unknown to the scribe. B1's careful consideration of what to change and what not to change demonstrates a deep commitment, where he uses all his resources to copy from a challenging and possibly at times illegible

exemplar, and skillfully tries to reconcile the old writings with the intuitions he has of his own language.

The second paper of Part 2, by Moessner, is a first quantitative analysis of the morphosyntax of Old English wills. It aims at complementing existing sociohistorical research, which has hypothesized that Old English wills are two-layered in that they combine an apparent but still emerging written mode with clear traces of their past origins of oral bequeathing (e.g. Danet & Bogoch 1994). In this respect, Old English wills can be expected to be less formal and less complex than their Present-day English (PDE) counterparts. Moessner's analysis confirms this expectation. Compared to PDE wills, Old English wills use significantly shorter sentences and have a lower average number of clauses per sentence. In addition, they hardly ever use the passive voice, whereas the passive accounts for a quarter of all verbal constructions in PDE wills. Another interesting feature of OE wills is the relatively low frequency of modal markers of obligation (in Old English *sceal* and the subjunctive); in contrast, *shall* is highly frequent in PDE wills. How this feature relates to the oral origins of wills is less clear, however, and definitely merits further research. In addition to this quantitative analysis of linguistic features, Moessner also draws attention to a number of other period-specific properties of the text structure of wills. Some of these further testify to the lower reliability of the written document, and reveal a different attitude of the Anglo-Saxon people to official documents. On the one hand, testators are aware of the future vulnerability of their wills and anticipate forgeries by including cursing formulae to discourage potential forgers. On the other hand, the common omission of witness lists in copies of wills suggests a different perception of what was considered a functional copy of an official document.

A different take on layers in text types is found in Nakayasu's paper. Nakayasu analyzes the two medieval text types 'fiction' and 'handbook' examining work by Chaucer (*Canterbury tales* and *Astrolabe*). The paper explores to what extent different types of distal and proximal deictic elements – i.e. spatial and temporal items, and their extensions (e.g. argumentative uses of *now* and *then*) – cluster and function as layers that add structure to the extra-linguistic experiences the author has in mind, and how these distal and proximal clusters are employed in different ways across the two text types. For some clusters, such as the proximal *here* and *now*, in combination with the use of the historical present, the distribution over the two text types confirms what we know from previous studies. While the fictional text type has an overall higher rate of distal deixis, due to the narration of past events, a shift from a distal past tense to a proximal present in storytelling was already purposefully used in Chaucer's texts to decrease the psychological distance between the event in the remote past and the contemporary hearer. Some less established usages of deictic clusters can be observed as well: these might have

been motivated by concerns of linguistically signaling distance between speaker and hearer or between hearer and topic. Examples include a connection between imperatives and distal elements, when the narrator Chaucer is giving instructions to his son in the *Astrolabe*, or a religious contrast between God (distal deixis) and the world he created (proximal deixis). Nakayasu also suggests that modal deixis (the distancing use of past tense modal verbs such as *should, would*, etc.) may be functionally linked to deictic elements from the spatial and temporal domains into a unified type of spatio-temporal deixis. Her quantitative evidence supports such a connection between spatial and temporal deixis in Chaucer that is stronger than is typically assumed.

The final contribution in this section (GATHER) discusses the link between language use and different attitudes by means of a syntactic analysis of the Reformation genre of congregational song. Essentially the genre consists of songs on a Christian topic, which were usually part of a liturgical service. They make use of metrical and rhymed English poetry in stanzaic form. Not only does Gather define this historical genre for the first time, but she also draws attention to its specific syntactic preferences, how these preferences change as the genre itself changes, and how they differ from secular poetry. Specifically, Gather analyzes the phenomenon of argument shifting, i.e. the 'improper' placement of object, complement, or obligatory adverbial as compared to S(Aux)V(X), the unmarked word order from the 16th century onwards. Gather shows that argument shifting is significantly more common in congregational song than in secular poetry. In addition, the genre itself changes, and with it the type of argument shifting. Early congregational songs were often based on psalms, and shared with them a tendency to move object pronouns to medial position, before the main verb (S(Aux)OV), in particular because of metrical constraints. Later on, hymns predominated. The most common type of argument shifting in hymns was that of argument fronting (as in the title *A riddle to myself I am*), which was mostly used for reasons of emphasis. What is interesting is that, regardless of the different types of syntactic-poetic behavior, the genre as a whole consistently behaved in a more archaic fashion than did secular poetry. On this basis, Gather argues that non-standard word order is not merely a matter of poetic license, but instead reveals a deliberate conservative attitude to the genre related to religious tradition, which adds a historical layer onto the genre. Congregational song also shows more archaic behavior than other religious genres; it shares this conservative stance with a genre like prayer, which also belongs more to the ritualized side of religion.

Lexis, morphology, and a changing society

Part 3 brings together a number of papers that show how a changing sociocultural situation may affect lexis and morphology.

The first paper (Ruano-García) provides an enlightening discussion of how different editions of Bartlett's 19th-century *Dictionary of Americanisms* reflects changes in the status of and attitudes towards American English, which was at the time developing into a distinct variant of English. The first edition of Bartlett's dictionary contained many words that were also in use in England. To some extent, Bartlett seems to have included these because they were only used regionally in England, but were in general use in the United States. Mostly, however, their inclusion reflects the mixed purpose of the dictionary: instead of focusing on words that were unique to the US or had significantly different meanings, the first edition in effect was more of a list of words that did not belong to standard British English, and were in use in (some parts of) the US. As one reviewer of the dictionary pointed out, this may have served "as a good purpose in separating the language of coarseness from that of elegance" (Anonymous 1848: 328). While therefore useful, it was not a dictionary of Americanisms in the strict sense. The situation changed with subsequent editions. A decade after its first publication, Bartlett added a more explicit list of conditions for a word to be considered an Americanism. Accordingly, many words that were also around in England were removed from the dictionary, and still more words were removed in the 1877 edition. Other words were added, when there was reason to believe they behaved significantly differently in the United States than they did in the UK. The overall effect of the consecutive editions was that, with the final edition, the dictionary had truly become a dictionary of Americanisms. The various editions, then, reflect some interesting sociolinguistic processes, which cannot be easily separated. To some extent they may reflect actual changes in the status of some words (which were in the process of becoming obsolete in the UK, but thriving in the US, for instance). More importantly, they show the emergence of an awareness of a distinctively American idiom, not only in Bartlett's editorial practices, but also in the reviewers' reactions.

Hotta, in the second paper, discusses the ebb and flow in the history of the variants of the prepositions *between, among*, and *amid*. His analysis reveals a remarkable parallelism in their development, which attests to their being a functionally coherent set. For all three, *-s*-forms (e.g. *amid[de]s*, based on the adverbial function of the genitive) became fashionable in ME, only to go out of fashion again and to be succeeded by *-st*-forms (*betwixt, amongst, amidst*), which became predominant in Early Modern English. From Late Modern English onwards, the *-st*-variant declined again, and the older forms, which had never been lost entirely, became most common again. Hotta draws attention to a variety of potential analogical associations

that may help explain this development. Apart from their similar function, it is argued that the *-st*-variant (even if the *-t* might have ultimately had a different origin) is an extension of the superlative ending on adjectives, with a similar intensifying function (*amid* "in the middle of" > *amidst* "right in the middle of").[2] However, such a distinction in emphasis, if it existed, was quickly bleached, and other factors took over the competition between the base form and its extended variants. Hotta suggests that an important reason for the decline of *-st*-forms may have been a change in attitude within the speech community. While these *-st*-forms were unobtrusive in Early Modern English, the frequency of consonant clusters ending in *-st* dropped dramatically in the course of Modern English because of the gradual loss of the second-person pronoun *thou*. In combination with the overall increased effort required to pronounce such consonant clusters, their increasing uncommonness may have made language users perceive them as archaic and stylistically marked, hence to be used with care.

Minkova, finally, provides a diachronic perspective on clipping or shortening after it being neglected for nearly five decades. Yet shortening is an important process in word formation, which today is responsible for 9%–15% of all new words. Minkova considers it premature to dismiss it as a phenomenon outside any linguistic rules (as for instance claimed by Don 2014), and shows how such a view fails to take into account the historical dynamics of clippings. Rather than operating in a random fashion, it is shown that the phonological and morphological factors in clipping act differently in different periods. With the exception of proper nouns, which are largely left out of the discussion, it is revealed that premodern English clearly favored foreclippings (type *mend < amend*, *print < imprint*). Largely, they adhere to the typical constraints of Germanic stress patterns: only unstressed onsets are lost, and the resulting form almost always starts with a full onset, i.e. a consonant (so *vangelist < evangelist*, not *angelist*). From the 16th century onwards, the rate of foreclippings decreases, and they also change in nature, now also including the deletion of more than a single syllable (e.g. *university > versity*) or the deletion of a morpheme bearing primary stress (*turnpike > pike*). More importantly, from the 17th century onwards, back-clipping takes over (*con < contra, mob < mobile*). The demotion of foreclippings is also particularly obvious in the restoration of the input left-edge material in relatively short-lived forms such as in †*sturb* (> *disturb*) or †*spittle* (> *hospital*). Such restorations also attest to a broader familiarity with the source language. Indeed, the massive influx of Romance loans, which goes through

---

2.   This constitutes an interesting case of morphological analogy across parts of speech, a topic that has not yet received much attention in the literature on analogy as far as we (the editors) know. A similar phenomenon may be observed in the occasional use in Dutch of *temiddenst* "in the middle of", *tussenst* "between".

a second wave in Early Modern English, ultimately leads to the abandonment of the typical Germanic word stress pattern in English. When stress was no longer a reliable indicator of word boundaries, phonological form became more important as an identifier; hence back-clippings became the rule, as they retained the phonological form of the beginning of the word. Today, attestations of foreclipping have turned out to carry more social stigma as being 'lazy', while PDE backclippings are simply 'cute', 'playful', or 'in-group', but lack the stigma of ignorance or carelessness.

Conclusion

The chapters collected here examine which structural features of lexis and text types are influenced by the sociocultural context. Together, they provide a varied window on the effect of historical versions of a dynamic society on the lexicon. Examining lexical and textual change in history from a sociocultural perspective teaches us a great deal – not just about the past, but it also makes us think about similar phenomena in the present, enhancing our knowledge about how universally human some of these phenomena are.

References

Anonymous. 1848. Critical notices: Dictionary of Americanisms: A glossary of words and phrases, usually regarded as peculiar to the United States. By John Russell Bartlett. *The American Whig Review* 8(3). 326–328.

Biber, Douglas & Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511814358

Biber, Douglas & Bethany Gray. 2013. Being specific about historical change: The influence of sub-register. *Journal of English Linguistics* 41(2). 104–134.

Danet, Brenda & Bryna Bogoch. 1994. Orality, literacy, and performativity in Anglo-Saxon wills. In John Gibbons (ed.), *Language and the law*, 100–135. London: Longman.

Don, Jan. 2014. *Morphological theory and the morphology of English*. Edinburgh: Edinburgh University Press.

Grant, Raymond J. S. 1989. *The B text of the Old English Bede: A linguistic commentary*. Amsterdam: Rodopi.

Hogg, Richard. 2009. Old English dialectology. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 395–416. Oxford: Blackwell.

Kestemont, Mike & Karina van Dalen-Oskam. 2010. Predicting the past: Memory based copyist and author discrimination in Medieval epics. In Toon Calders, Karl Tuyls & Mykola Pechenizkiy (eds.), *Proceedings of the twenty-first Benelux Conference on Artificial Intelligence (BNAIC 2009)*, 121–128. Eindhoven: TU/e Printservice.

Kestemont, Mike, Sara Moens & Jeroen Deploige. 2015. Collaborative authorship in the twelfth century: A stylometric study of Hildegard of Bingen and Guibert of Gembloux. *Digital Scholarship in the Humanities* 30(2). 199–224. doi:10.1093/llc/fqto63

Kouchakia, Maryam & Francesca Ginob. 2016. Memories of unethical actions become obfuscated over time. *PNAS* 113(22). 6166–6171. doi:10.1073/pnas.1523586113

Laing, Margaret & Roger Lass. 2009. Early Middle English dialectology: Problems and prospects. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 417–451. Oxford: Blackwell.

Laing, Margaret, Angus MacIntosh & Michael L. Samuels. 1989. *Middle English dialectology: Essays on some principles and problems*. Aberdeen: Aberdeen University Press.

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

Nichols, Stephen. 1997. Why material philology? Some thoughts. In Helmut Tervooren & Horst Wenzel (eds.), *Philologie als Textwissenschaft: Alte und Neue Horizonte* (Sonderheft Zeitschrift für deutsche Philologie 116), 10–30.

Stein, Dieter & Ingrid Tieken-Boon van Ostade (eds.). 1993. *Towards a standard English: 1600–1800*, vol. 12. Berlin: Mouton de Gruyter. doi:10.1515/9783110864281

Conspicuous lexical choice in past societies

Old English *ead* in Anglo-Saxon given names

A comparative approach to Anglo-Saxon anthroponomy

Olga Khallieva Boiché
Independent Scholar

The Anglo-Saxon name-element *ead-* was the one of the most common themes used to form dithematic personal names. The overall goal of this paper is to examine the reason for this abundance in anthroponyms such as *Eadbald*, *Eadberht*, *Eadred*. First, the paper addresses the questions of which meaning of Old English *ead* (wealth, prosperity, happiness, bliss) is present in personal names, and whether there is a hidden message in these names. I look for the ancestral meaning of *ead* by examining the Old English poetic corpus and then establish lexical and anthroponomical parallels within continental Germanic and other Indo-European traditions: namely Old Greek and Old Slavic. The paper suggests that names formed with Germanic *aud- (etymon of OE *ead-*), Greek *plouto-*, and Slavic *žir-* referred to riches and abundance and were originally used within the inferior stratum of society. Originally, they would have been monothematic and reflected the parents' desire that their progeny would have a prosperous life. Dithematic anthroponyms, then, containing the *ead-* element (*Eadric*, *Eadwald*, *Edmund*) emerged during the Heroic Age; they were a later interpretation of ancient poetic epithets describing the lord as a guardian of his people and his land.

Keywords: personal names, Anglo-Saxon, Germanic, Old Russian, Old Slavic

1.   Introduction

The (Proto-)Germanic name-element *auda, corresponding to the OE noun *ead*, is one of the best preserved themes in modern given names inherited from Old Germanic times. From earliest times, Germanic speakers have demonstrated their preference for this name-element, as can be seen from the fact that *Otto*, *Edward*, *Edmund*, *Edwin*, and *Edith* became frequently used cosmopolitan names.

The definition of the OE lexical entry *ead* provided in *Bosworth & Toller's Anglo-Saxon dictionary* (*BT*) is "possession, riches, prosperity, happiness, bliss" (*BT*, s.v. *ead*). The question, then, is which semantic notions, "wealth" and "prosperity" or "happiness" and "bliss", are present in the personal names? A further question is whether there is any coded message is conveyed by these names and whether this message could indicate any social affiliation?[1]

These questions are addressed below by following the model proposed by Schramm (1957) and developed by Schmitt (1973), according to which dithematic personal names were the expression of poetic language.[2] On this view, dithematic names in the Old Germanic languages corresponded originally to fixed poetic epithets or formulaic expressions that existed at the period of name creation (Schramm 1957: 58–59). Dithematic personal names are considered as primary formations not only if they represent an intelligible syntagm but if this syntagm is attested in the poetic speech of the source or of a related language or if it can be reconstructed with the help of comparative grammar. The majority of dithematic given names cannot be easily associated with any intelligible syntagm. These are the secondary formations which are the product of parents' or close relatives' recycling of given names, when a person is linked to his/her family members by sharing with them either the same name-element (the principle of variation) or the same initial letter (the principle of alliteration). Therefore, as Schramm (1957: 145–147) and Schmitt (1973: 46) put it, the task of modern anthroponomy is to look for the primary dithematic anthroponyms and find corresponding poetic epithets rather than to discuss the intelligibility of Old Germanic names.

This point of view is also shared by modern scholars in Anglo-Saxon (Gneuss 1991: 51; Insley 2007: 9), Germanic (Greule 1996), and Classical studies (Masson et al. 1990: 1, vii–xii).

Another perspective on Old English personal names is presented by Fran Colman, who states that "names have reference, but not sense" (1992: 14).[3] She separates etymology and meaning:

---

1.   In name-giving, parents may code a message that does not reflect the etymology of the name but some hidden reference. For instance, a child may be named after a person the parent(s) admire(s). As will be seen, *ead*-names were given to procure a bright future for the newborn; originally, these names were not used by the elite.

2.   Dithematic personal names are names made up of two stems; for instance *Æðelstan* comprises two elements: OE *æðel* "noble" and OE *stan* "stone".

3.   In her recent book Fran Colman reiterates this position (Colman 2014: 1–18).

> [T]o claim an etymological 'origin' of a proper name-element is not to claim that the element retained the range of denotations available to the cognate common word … . Acquiring a solely referential function, the name-element is freer to go its own way, with respect to phonological (and/or orthographic) variation.
>
> (Colman 1992: 12–15)

I cannot fully agree with this, because it presupposes that anthroponymy is an in-variable system. Rather, anthroponomy is productive and constantly changing. On the one hand, there is a general tendency for names and name-elements to lose their meaning in the course of time, in the process acquiring only referential function. This is a process that takes centuries, and the Anglo-Saxons "had advanced about half-way towards the present-day indifference to the signification of the names" (Redin 1919: 184). On the other hand, the list of proper names in a culture (the onomasticon) is constantly updated with newly coined meaningful names. This can best be illustrated with the example of modern American given names. While the great majority of them have lost their original meaning, new meaningful names are coined in the modern period: *Bliss*, *Blossom*, *Hunter*, *Chase*, *Cherry*, *Daisy*, *Delight*, *Desyre*, *Diamond*, *Duke*, *Scarlet*, *Skipper*, *Speed*, *Stone*.[4] They are so transparent that there is no need to indicate to which sex they should be attributed. This situation holds for other modern cultures as well.[5] In sum, along with the global tendency for names to acquire a solely referential function, there is a constant universal human need to create meaningful names.

The paper proceeds along two axes. First, it looks for the ancestral, original meaning of the (Proto-)Germanic noun *auda (OE *ead*) by examining the Old English poetic corpus (*ASPR*, *CASPR*). Second, it establishes lexical and anthro-ponomical parallels within continental Germanic and other Indo-European traditions: namely Old Greek and Old Slavic. The Anglo-Saxon and continental Germanic anthroponomy is thus treated not as an isolated phenomenon but as a part of the common Indo-European tradition.

The paper is based on the following data. All the poetic occurrences of *ead* and its synonyms are identified with the help of *A concordance to the Anglo-Saxon poetic records* (*CASPR*). Anglo-Saxon given names are drawn from the original core of *Durham Liber Vitae* (*DLV*) edited by David and Lynda Rollason in three volumes with linguistic commentary (Insley & Rollason 2007; Insley et al. 2007) and from

---

4. The basic principle of transparency or intelligibility is the availability of the corresponding word in the contemporary language, so "it can be assumed that the [bearers] themselves could [or can] easily associate them with the corresponding words in the living … vocabulary" (Redin 1919: XLII).

5. The author can assert this for the modern East Slavic and Turkic societies.

the *PASE* database (*Prosopography of Anglo-Saxon England*), taking into account only persons born before 900. The continental Germanic onomastic data are based on the list of the persons attested as living before 500, extracted from the work of Reichert (1987). Compilations of Slavic personal names recorded before 900 and Russian personal names recorded before 1400 are used to exemplify the Slavic material (Khallieva Boiché 2015: Annexes 4, 5). The Greek examples are drawn from *A lexicon of Greek personal names* (Fraser & Matthews 1987–2010).

2.   Old English *ead-* and *wel-* and their cognates in Germanic poetry and anthroponomy

In the earliest Germanic languages, *auda denoted either riches or happiness (Lühr 2017: 1235–1239). Pokorny's *Indogermanisches Etymologisches Wörterbuch* establishes the initial meaning of (Proto-)Germanic *auda-* as "happiness, possession, wealth" (Pokorny 1959–1969: I, 76).[6] Lehmann, on the basis of Wulfila's translation of Greek *makarios* "blessed" as Gothic *audags* (Lehmann 1986: A223), notes that there are no clear cognates of (Proto-)Germanic *auda in other Indo-European languages. He remarks that the Germanic forms reflect belief in fortune determined by fate. Table 1 presents *auda- cognates in the Germanic languages, drawing on Lehmann and Lühr.

As can be observed, the Germanic languages preserve this duality of meaning: "possessions" and "happiness". Old Saxon (OS) uses this root mostly to refer to material goods and wealth, as does Old High German (OHG). Gothic uses Greek *makarios* "blessed", while wealth is represented by the word *gabigs*. Ochs (1921: 102–112) explains this seeming specialization as a peculiarity of Wulfila's style. ON (Old Norse) cognates show only the meaning "wealth" and "given by fate".

In defining the (Proto-)Germanic adjective *audag, which corresponds to Old English (OE) *eadig*, both the *Etymologisches Wörterbuch des Althochdeutschen* (Lühr 2017: 1239) and Lehmann (1986: A223) refer to Ochs, who states that *audag means "favored by fate" (Ochs 1921: 103–104). Lehmann adds Weisweiler's definition "gifted through fate" (Weisweiler 1923: 318–319). However, the presence of the notion FATE does not find justification in (most of) the cognates of *ead(-)*: the notion of fate is present only in ON. At the same time, the other Germanic languages do reveal a connotation of the supernatural, of something beyond our control and

---

6.   Kroonen (2013: 40–41) reconstructs the (Proto-)Germanic meaning of *auda- "riches". However, he does not include in his list Old Saxon *od* and Old English *ead*, both of which denote not only "riches" but also "prosperity, happiness" (Sehrt 1966: 421; *BT*, s.v. *ead*).

Table 1. *Auda- cognates in the Germanic languages

Old English (OE)		Old Saxon (OS)		Old High German (OHG)		Old Norse (ON)		Gothic	
Ead	riches prosperity happiness bliss	*ôd*	property prosperity happiness						
Eadig *Eadeg*	happy blessed prosperous rich	*ôdag*	rich in possessions	*ôdeg, ôtag*	rich, opulent	*auðogr*	rich	*audags*	blessed
Eaden	given, granted	*ôdan*	granted born			*auðenn*	fated		
Eadwela	riches happiness blessedness	*ôd-welo*	inherited property riches	*ôt-uuála*	riches possessions				
				ôt-púdela	treasurer				

deprived of free will. This supernatural connotation may explain Wulfila's choice in translating the Greek *makarios* "blessed" as Gothic *audags*. In Greek, *makarios* is already used by Homer to mean happy or joyful when he refers to gods and sometimes men who are carefree and favored by gods (Chantraine 1968–1977: 3, 659). It therefore seems appropriate to adjust Ochs's and Weisweiler's definition of *audag to "favored or granted by the divine powers" and, accordingly, to say that OE *ead* denotes a divine gift. In (Proto-)Germanic, this divine gift probably referred to concrete, material wealth. The abstract spiritual meaning developed later: with the advent of Christianity this divine gift takes the form of blessedness, so that the OHG adjective *otag* translates Latin *beatus* "happy".

2.1   *Ead-* in Old English poetry

In search of the ancestral meaning of *ead*, I examined poetic sources in the Old Germanic languages. In the Anglo-Saxon poetic corpus, 31 out of 132 occurrences of *ead* and its derivatives refer to wealth and prosperity (*CASPR*, s.v. *ead*). Taking into account the share of religious poems in the Anglo-Saxon poetic corpus, this number is quite high and revealing. These are:

ead "riches, prosperity, wealth": *Genesis* 1602, 1891; *Daniel* 671; *The riming poem* 31; *Exodus* 186, 339; *The ruin* 36; *Riddle* 26, 23; *Guthlac* 1192; *Precepts* 89;

> *eadig* "rich, prosperous": *Riddle* 84, 28; *The rune poem* 76; *Maxims* 1, 156; *The order of the world* 97; *Genesis* 1878, 2148; *Christ* 1496; *Exodus*, 29; *The judgment day* 2, 163; *Beowulf* 2470; *The fortunes of men* 60; *The metrical charms* 2, 40;
>
> *eadgifa* "giver of prosperity or happiness": *Andreas* 74, 451;
>
> *eadwela* "prosperity, abundant prosperity, happiness, wealth": *Phoenix*, 251; *Andreas* 808; *Elene* 1316; *Guthlac* 1091, 1118; *The fortunes of men* 67; *Phoenix* 586.

An examination of these examples leads to the conclusion that the Germanic adjective *audag does not have the general meaning "given by the divine powers", but refers more specifically to "material goods given by the divine powers". In other words, the notions of "wealth" and "prosperity" attested in the North and West Germanic cognates of *audag indicate that *audag was initially associated with a material gift. On Ochs's view, however, the basic meaning of *audag, namely "given by fate" (1921: 103–104), operates as a shortcut for the abstract meaning HAPPY (if one is granted something by fate, then one is happy). I argue, though, that in cases such as this, when the concrete meaning is attested in several languages along with the abstract meaning, it can be safely assumed that the concrete meaning existed from the very beginning and that the abstract meaning was derived later. This is in line with principles of semantic change. As Vajda explains: "There are universal tendencies in the directionality of polysemy … . Directionality in polysemy seems to be logically motivated: concrete meanings give rise to abstract ones (sharp knife -> sharp mind)" (Vajda 2001). Therefore, it is argued here that the semantic characterization "material goods given by the divine powers" is the original one and that the reference to happiness and blessedness is a later development.

This question is crucial for the understanding of personal names containing the element *ead-*: to which notion does the anthroponomical element *ead-* refer in the personal names – to "luck and happiness" or to "wealth and prosperity"? Ochs (1921: 103–104) asserts that in the anthroponyms we observe the ground meaning of "gifted by destiny". I argue that the ground meaning was "wealth given by the supernatural forces or gods" and that this meaning of material wealth as opposed to spiritual wealth is preserved in the Germanic given names.

2.2   OE *wela*, Old Saxon *welo*, OHG *uuóla* "riches, prosperity" in poetry and anthroponomy

A characteristic of OE and more generally Old Germanic poetry is the great availability of synonymous epithets. According to Schramm (1957: 145–147) and Schmitt (1973: 32–36), Old Germanic compound names reflect this particularity

of the poetic language. The best example is the quantity of terms referring to battle which are both poetic words and name-elements: OE *hild, gūð, heaðu, beadu* "war, battle".

In Anglo-Saxon poetry, the elements *ead* and *wela* are attested as synonyms meaning "wealth, prosperity and riches". Whereas *ead* and its derivatives with the meaning of riches are encountered 31 times in the Anglo-Saxon poetic corpus, *wela* and its derivative *welig* "wealthy, rich, opulent" appear 51 times.

Cognates of OE *wela* also occur in continental Germanc languages. In OS, *welo* "riches, property, possessions" and a derivative adjective *welag* "rich" are widely used as well (Sehrt 1966: 650). In *Heliand* and *Genesis* there are 17 occurrences of this lexeme.

> *welo*: *Heliand* 871, 1325, 3143, 1840, 2120, 4923, 1098, 1330, 2137, 2643, 2113, 1677, 3293, 3773, 2488, 1023; *Genesis* 262.

In OHG scholarly literature, the cognate of OE *wela* is attested in the nouns *uuóla, uuolo* "happiness, happy estate; delight, enjoyment; money, wealth" and *uuëralt-uuolo* "money, riches" (Sievers 1872; Piper 1883; Kelle 1889: 1–154; Sehrt & Legner 1955; Sehrt 1962).

In sum, the elements *ead* and *wela*, and their cognates, are attested in OE, OS, and OHG literary works as synonyms. Importantly, these elements are also used in Germanic anthroponyms. It seems likely, therefore, that both themes are also used as synonyms in the Germanic personal names.

While we do not find name-elements among the Germanic personal names that could be synonymous with OE *ead* in its meaning "happiness, luck given by supernatural forces", we do find synonyms of *ead* in personal names referring to material wealth and prosperity. Thus, of the several synonyms in Anglo-Saxon poetry of *ead* with the meaning of "wealth, prosperity and riches", i.e. *æht, gestreon, sped, wela*, the first three elements are not represented in anthroponomy, but the fourth, *wela*, is also a name-element among the Germanic personal names, as is shown by Förstemann (1900: 1551–1555), who presents the following list of continental names of the 6th to the 10th century:

> *Weli, Welo, Wela, Weling, Weliga, Welaperht, Welipurc,*
> *Welatrud, Welafrid, Welagrim, Welcont, Welheid, Wellehart,*
> *Welharius, Velaramnus, Weliman, Welamot, Welamunt, Welarat,*
> *Welarata, Welrih, Welisind, Wealo, Weala, Veland, Vialber,*
> *Wialbret, Wielburg, Wealdrud, Wealfrid, Wiolgart, Wiolicus,*
> *Wielind, Wielioz, Wielachoma, Wielrat, Wealsind, Weololf,*
> *Welaspurc, Welisung.*

2.3   Weland

With regard to the anthroponomical data, Förstemann (1900) is hesitant in translating the name-element *wela*, noting that it is difficult to decide from which stem it originates: from OE *wilian – wilwan* "to roll, to join" or from ON *vêl* "art, profession, craft".

Consider the personal name *Weland*, of which Förstemann (1900) lists around 20 occurrences. The earliest mentions are *Veland*, attested on a 7th-century stone inscription in Mainz, and *Weland* of the 7th-century Frisia (Reichert 1987: I, 770–771). In the British Isles, this name is associated with the legendary smith Weland, but OE preserved only scanty references to Weland's story. In ON sources, Völund (OE Weland) is a main character of *Völundarkviða* (*The lay of Völund*), a poem in the *Poetic Edda*[7] (Jónsson 1949; Adams Bellows 1936). It relates how three bachelor brothers, Völund (OE Weland), Egil, and Slagfiðr, venturing out one night, come across three valkyries spinning flax by the river. After the young maidens have removed their swan cloaks, the brothers seize them, bring them back to their home, and take them as their wives. After nine years, the valkyries fly away, deserting their husbands. The heartbroken Völund stays at home and constructs replicas of his wife's gold ring, gradually developing masterful smithing skills. Soon after, King Niðuð, greedy for gold, kidnaps Völund and imprisons him on an island. There Völund, hamstrung, is forced to forge items for the king. Eventually he finds means of revenge. He kills Niðuð's sons, impregnates his daughter, and then flies away using wings he made.

The etymology and meaning of the name *Weland* are not clear. Grimm was the first to link up the name with ON *vêl* "art, profession, craft". Other suggested meanings are: ON *vel* "cunning", *vela* "to outwit, to cheat", or *wel-nand* "brave in battle" (Rosenfeld 1969: 53–62; Nedoma 1988: 59–67). In my opinion, the name *Weland* includes the component *wela* "riches, property, possessions" as its first part. Support for this interpretation may come from *Völundarkviða*, which as Talyor (1963) suggests may be interpreted as an account of a ritual regeneration.

> The plot outline, expressed in mythic terms, is this: a man has been visited by a swan-maiden for eight years, but she leaves him in the ninth. During this period … **he is prosperous and productive**. At the end of the ninth year, however, he is attacked by a winter or death force (Nithuthr) who robs him of his strength and sets him on an island to produce weapons. The captive regains his power through a sexual act which gives him the ability to raise himself (as spring vegetation) while his captor loses his power simultaneously.   (Taylor 1963: 234; emphasis mine)

---

7.   The Poetic Eddas were written down from 1000 to 1300 CE. Henry Adams Bellows' translation of *Völundarkviða* [The Lay of Völund] is available on line: http://www.sacred-texts.com/neu//poe/index.htm.

This interpretation may throw light on the semantics of the hero's name. In Taylor's words, Weland is one who is "prosperous and productive" (Taylor 1963: 234). Another link between Weland and the world of prosperity and abundance is his association with the elves. Twice in the poem we find this reference: *alfa ljóði* "the master of elves", *vísi alfa* "greatest of elves" (*Völundarkviða* 10, 32). As Ellis (1943) has pointed out, the elves' cult can be associated with abundance and prosperity. Furthermore, the Scandinavian god Frey, associated with farming, abundance, and fertility, is said to be the master of the realm of the elves in another *Poetic Edda* poem, *Grímnismál* (Jónsson 1949: 5; Adams Bellows 1936: 5). In the medieval Scandinavian traditions, there is a strong correlation between elves, the dead, mounds, and the fertility of the earth. This affinity was also illustrated and explained by de Vries (1956 [1935–1937]: I, 256–260) and Lecouteux (1997).

There is also a view that Völund (OE Weland) is the survival of the divine figure. Thus, according to Motz (1986: 51–52), the core part of *Völundarkviða* resembles the mythological themes of Uralo-Altaic and Paleo-Siberian peoples. Recurrently in this mythology, the son of the sky god descends to earth, usually in the shape of a duck, and here (on earth), he performs the office of a healer and a shaman. Sometimes, however, he suffers injury at the hands of men and takes vengeance.

In the light of these tales, we may consider *Weland* as the name of a god rather than that of a smith, which may explain why *Weland* is not associated with a craft.

2.4   Plouton

Confirmation of the association of *Weland* with "riches, property, possession" can be found in Old Greek. Indeed, a strong parallel to the name *Weland* exists in Greek mythology, namely *Plouton*, the name for the god of the underworld. This name derives from Greek noun *ploutos* "riches, opulence, abundance in possessions" (Chantraine 1968–1977: 3, 918) and denotes "the one who has wealth". The fact that both *Weland* and *Ploutos* are built on the lexeme referring to wealth and prosperity may not be fortuitous.

According to Nilsson, the name *Plouton* was attached to the god of the underworld *Hades* only at a later stage in his history (1941: 111–112, 296, 424, 442). Originally *Ploutos* was the name of the son of the goddess Demeter, the goddess of the fertility of the earth and of the seasons. Homer writes that Demeter fell in love with Iasion and lay with him on the thrice tilled field. Hesiod continues this story, specifying that as a consequence of this intercourse Demeter had a son named *Ploutos* "the stock of grains", the god of riches. Nilsson (1941: 111, 112) further points out that in the Greek pantheon there was no place for Ploutos near Demeter's daughter Kore, so the cult of Ploutos gradually died out. It underwent a sort of rebirth in the 4th century BCE when his name merged with that of Hades.

2.5   Greek *plouto-* anthroponyms

Like Germanic **auda*, the Greek noun *ploutos* forms numerous anthroponyms. In *A lexicon of Greek personal names*, names containing the element *plouto-* occurred widely in all the Greek provinces (Fraser & Matthews 1987–2010). The following names are recorded:

> *Plous, Plousia, Plousis, Ploutarxa, Ploutas, Ploutarxe, Ploutarxis, Ploutarxises, Ploutarxos, Plouteros, Ploutianos, Ploutias, Ploutiases, Ploutina, Ploutinas, Ploutinos, Ploutis, Ploution, Ploutixa, Ploutogenes, Ploutogeneia, Ploutos.*

Among all these names there are only two dithematic ones: *Ploutarxos* and *Ploutogenes*; the other ones are monothematic and monothematic hypocoristic formations.[8]

In sum, the name *Weland* finds a parallel in the Greek *Plouton* and *Ploutos* of Greek mythology and anthroponomy. This parallel reinforces the hypothesis that the name-element *wela-* coincides with the OE noun *wela* "wealth, riches, prosperity".

2.6   Germanic anthroponyms containing the *wel-* element: The OE compound *eadwela* and its Germanic cognates

I now consider the other Germanic anthroponyms containing the element *wela-/wel-*. Among the earliest Germanic names listed by Reichert (1987: I, 771), there are several with the stem *wēl* (which differs from the stem *wel* that is included in the name of the Germanic seer Veleda, mentioned by Tacitus): *Vel* (the Gepid of the 6th century), *Veldes* (Italy), *Vellango* (Cologne).

In the British Isles, there are no masculine names with the element *wela-/wel-*. On the other hand, Okasha (2011: 51, 67) cites two feminine names attested before the end of the 9th century: *Weale* and *Wealenburg*. Okasha interprets the name *Weale* as related to OE *wealh* "Celt, slave", and she considers the form *Wealenburg* obscure. However, it can be argued that both names bear the element *wel-*, corresponding to the OE masculine noun *wela/weala/weola* "riches, prosperity". Morphologically, the form **weale* is the feminine form of the masculine noun *weala* (weak declension, *-an* stem). Morphologically, *Weale* looks like other feminine OE personal names: *Bote* (m. *Bota*), *Nunne* (m. *Nunna*), *Ætte* (m. *Atta*).

---

8.   Monothematic personal names are names that are made up of one stem; for instance *Offa, Demma, Eada.*

The second name, *Wealenburg*, can be argued to include the OE noun *weala* in the genitive plural: *wealena* (with dropping of the final *-a*); the fact that *weala* is often employed in plural corroborates this hypothesis. This personal name was probably coined in the 9th century (*Wealenburg* is attested in *S* 277, dated 833)[9] by analogy with late continental forms, where the first element is used in Genitive: *Gotesschalk* "God's slave", *Gotesman* "God's man". The creation of the name *Wealenburg* may also have been prompted by place names, whereby the ancient anthroponomical theme *burg-* was reinterpreted in accordance with the existing OE word *burg* "town, fortress". In this case, the personal name *Wealenburg* would designate "a town of wealth". Within modern British topography, there is a synonymous appellation: Eddisbury in the county of Cheshire. *The Anglo-Saxon chronicle ad a 913* tells us that "*Æðelflǽd Myrcna hlǽfdige ða burh getimbrede æt Eádes byrig*" [Æthelfled, lady of the Mercians, built a fortress at **Eddesbury**] (*BT*: *Eades burh*). Both the appellations *Wealenburg* and *Eades burh* are completely synonymous if we consider that *weala* and *ead* mean "wealth" in OE.

Continental dithematic personal names containing the *wel-* theme are mentioned by Förstemann (1900: 185, 203): *Wallod* (7th century), *Odewalo* and *Oduvalla* (9th century). The specific feature of these names is that they correspond to the OE nominal compound *eadwela* "happy weal, riches", to OS *ôd-welo*, or to OHG *ôt-uuála* "inherited possession, riches". There are seven occurrences of this compound in OE poetry.

Bradley (1982) translates *eadwela* either with the unique words *happiness* or *prosperity* or with such redundant collocations as *happy prosperity* or *abundant prosperity* (*Phoenix* 251; *Andreas* 808; *Elene* 1315–1316; *Guthlac* 1090b–1091; *Guthlac* 1117–1118; *The fortune of men* 67; *Phoenix* 586). Such translations allow us to consider *eadwela* as a copulative compound,[10] where two lexemes with a close meaning are joined together to reinforce the meaning of the compound, and the corresponding names *Odewalo* and *Wallod*, as examples of primary formations (see Section 1), were obviously brought into existence following the copulative compound *eadwela*.

---

9. 'S' stands for *Sawyer's catalog of Anglo-Saxon charters* (see Data sources).

10. Copulative compounds (or dvandva) are those, where two semantic units can be connected by an *and*, for instance, *bittersweet*, *spacetime* (Schmid 2016: 126).

2.7   Anthroponomical implications of the poetic use of *ead*

A comparison of the use of *ead* with that of its synonyms in OE poetry leaves the impression that the meaning of *ead* is more abstract than the meaning of *wela*.[11] When *ead* is used in contrast with *earm* "poor", it denotes the rich (*Riddle* 84, 28; *The rune poem* 76; *Christ* 1496; *Exodus*, 29; *The judgment day* 2, 163; *The metrical charms* 2, 40). When *ead* is used with *wela* and *æht* (i.e. within the same line/sentence in the OE poetic corpus), it denotes the prosperous, while *wela* and *æht* in this case denote the material goods. Thus, the use of *wela* and *æht* in these cases helps to delineate the meaning of *ead* (*Daniel* 671; *The ruin* 36; *Genesis* 1878; *Maxims* 1, 156). It would seem, then, that *ead* had the meaning of wealth given by gods or destiny: it could denote both material and spiritual wealth, but with the connotation that it was given by the divine powers. This connotation is absent in the synonyms *wela* and *æht*.

If *ead* is of common Germanic stock, *wela* is attested only in West Germanic. The same can be said of the distribution of the personal names containing *ead-* and *wela-* elements. *Wela-* names are not known outside the West Germanic area, while *ead-* names were also in use in East and North Germanic. Thus, it is known that Theoderic the Great had Odvulf as an ancestor. Even if the historical existence of this ancestor has not been proven, the name itself must have been known to the Goths. As for North Germanic, Gering (1903: 1404) affirms that the noun *aud* "wealth" is conserved among the Scandinavians only in their personal names.

Two conclusions follow from this. First, Germanic *wela-* names should indeed refer to OE *wela* "wealth" and its cognates and not to the Germanic *wela* "battle", OE *wilian/wilwan* "to roll, to join", or ON *vêl* "profession, craft". Second, the *ead-* names as common Germanic ones are older than *wela-* names; in this respect, it may have been the case that *wela-* names were coined as a variation to *ead-* names and that the origin of this variation was the form *Odewalo/Wallod*.

The numerous personal dithematic names containing the elements *ead-* and *wela-* do not have an exact correspondence in the poetic epithets. This leads us to suppose that originally the names containing the element *ead-* were monothematic or not compounded.

This idea finds further support in the Greek data. Among the numerous name-forms containing the theme *plouto*, synonymous with *ead*, there are only two dithematic forms: *Ploutarxos* and *Ploutogenes*. An additional dithematic name is *Ploutakandros* (Ploutakandros Androneikou) mentioned by Bechtel (1917: 372). Bechtel notes that the name of origin had to be *Ploutax* and that the

---

11. For the detailed analysis of the poetic use of *ead*, see Khallieva Boiché (2015: 100–101).

form *Ploutakandros* results from adding the element of the father's name, in this case *andros*. It should be added that the name *Ploutax* corresponds to the Greek noun *ploutax* "a rich person" with a strong pejorative connotation (Chantraine 1968–1990: 3, 918).

This Greek form *Ploutax* enables us to look at its equivalent in Germanic anthroponymy, namely personal names with the element *Aud-*, from a new angle. Thus, Ochs (1921: 109) links the name of the 6th-century Suebian king *Audeca* with the ON adjective *auðogr* "rich". This form finds a close equivalent also in OE *eadig* and OHG *ôdeg* "prosperous, rich". To this can be added such early attested names as *Adico* (Roman soldier of the 3rd century), *Adeg*, and *Adic* (both of the 6th century; Reichert 1987: I, 10–11). Among other early simplex names containing *ead-*, Reichert identifies *Otto* and *Audil* (Franks of the 6th century), as well as four individuals, including one woman, named *Aud* (3rd–6th century).

Förstemann lists uncompounded occurrences as *Odi, Oti, Ode, Audo, Auto, Oudo, Odda, Odho, Otho, Oto, Otto, Ouda* (1900: 185–189). In Anglo-Saxon England, occurrences of monothematic *ead*-name before the end of the 9th century are represented in the *Prosopography of Anglo-Saxon England* by four individuals named *Eada* and three named *Eoda* (*PASE*). The original core of *Liber Vitae Dunlemensis* (first half of the 9th century) reveals 273 *ead-* names, 23 of which are uncompounded (Insley & Rollason 2007: 165–187):

In this section, I have defined Germanic anthroponomical themes that carry the notion of worldly riches and wealth. This notion does not suit the anthroponomical stock of a Heroic society with its warlike values very well.[12] To examine the onomastic practice within another Indo-European tradition, I turn to Old Slavic anthroponomy.

---

12. Germanic poets made "use of a large stock of distinctly poetic words" (Gneuss 1991: 47). The same stock was used in name-making. The great number of Germanic name-elements are synonymous poetic terms dealing with war (Gneuss 1991: 47–52; Schmitt 1973: 32–35; Schramm 1957: 144–147; Mallory & Adams 1997: s.v. *fame*; *name*; *poetry*). As Woolf put it almost hundred years ago, "in Germanic, battle had become a preoccupation of members of all social levels, and the distinctions between levels had become increasingly diffuse. This phenomenon was called by Chadwick the Heroic Age, and it reflects perfectly such militarization of the Germanic peoples" (Woolf 1939: 120).

3.   Slavic *žir*

3.1   The lexeme for "riches" in Slavic languages

The common Slavic root expressing the notion of wealth is *bog*. The present-day Slavic languages have conserved only derivatives of this word: Russian *bogatyj* "rich", *bogatstvo* "wealth, riches". The term *bogŭ* initially (during the pre-Christian period) signified both wealth and its giver. Later on, the Slavs assigned to this term the general meaning of "god" (Jakobson 1950: II, 1025). This process can be compared with that in Old English: the OE poetic compound *eadfruma* applies to God and is translated as "author of happiness" (*BT*, s.v.) or "giver of prosperity" (Clark Hall, s.v.; *Andreas* 1292, *Christ* 531–532).

Among the Old Slavic anthroponyms of my corpus, which includes West and South Slavic names encountered before the 9th century and East Slavic names encountered before the 14th century, there are very few names containing the theme *bog-* (Khallieva Boiché 2015: 391–454). I found only seven occurrences of such names: *Bogdan* (two), *Bogša* (two), *Boguslav* (two), *Boželisa* (one). In the name *Bogdan* "given by god" – probably a calque of Latin *Deodatus* – *bog* clearly refers to "god". The interpretation of the other three forms is uncertain, although Miklosich (1927: 11) states that all Slavic names containing *bogŭ* refer to god.

3.2   Slavic anthroponyms containing the *žir-* element

The visible absence of semantic elements denoting PROSPERITY and WEALTH in Slavic anthroponyms is challenging and enigmatic. At the same time, numerous occurrences of the name-element *žir-* are found in the corpus of Old Slavic personal names, which requires a thorough analysis. The name-element *žir-* is a component of such current Old Russian names as *Domažir*, *Žirjata*, *Žiroslav*. In what follows, all the occurrences of such names in my corpus are presented (see Khallieva Boiché 2015: 429–454).

Slavic Personal *žir-* names encountered before the 9th century are: *Histoižir* (*Histoisir*), *Našežir* (*Nazisir*), *Nežir* (*Nesir*), *Semižir* (*Semigir*), *Svjatžir* (*Suitschir*), *Žirna* (*Zirna*), *Žirnev* (*Zirneu*). In this early stage, *žir* names were especially common in Pannonia and Carinthia. In the Pannonian charter of the 11th century, there are three more uncompounded *žir-* names: *Girnus* (*Žirnus*), *Zirno* (*Žirno*), *Zirnugo* (*Žirnogo*) (*CDRCDS* 1967). These names have disappeared from the later sources: they are not found among the Bulgarian Slavic names. Here are the *žir-* names encountered among the Eastern Slavs before the end of the 14th century: *Dobrožir*, *Domažir*, *Mutižir*, *Nažir*, *Negožir*, *Sežir*, *Stežir*, *Xot'žer*, *Žirjata*, *Žirjač*, *Žirko/a*, *Žiročka*, *Žironežk*, *Žiroša*, *Žiroslav*, *Žirovit*.[13]

---

13.   For detailed information on each name, see Khallieva Boiché (2015).

In the territory of the Western Slavs, medieval sources report only one name containing the element *žir-*: it is also a monothematic name, *Žir* (*Syre*), recorded among the Sorabs in 1317 (Schlimpert 1978: 163). To West Slavic can be added Old Czech names: *Kromežir, Mojžir, Nedažir, Neužir, Ranožir, Svojžir, Vratižir, Vrtežir* (Svoboda 1964: 93).

3.3   The lexeme *žir* in the Slavic languages

Apart from being a name-element, *žir* is a common Slavic word comprising the following cognates (Šanskij 1962–: 9, 293):

> Russian, Bulgarian *žir* "fat"
> Ukrainian, Belorusian, *žyr* "feeding"; *žyrovanne* "abundance in food"
> Polish *żyr* "animals' food, food"
> Czech *žir* "animals fattening, food"
> Serbocroatian, Slovenian *žir* "gland" (as food for pigs)

The question arises as to why a theme with such a concrete grounded value is so common in the earliest Slavic names. Did another lexical meaning of the word *žir* exist? Another meaning is found in the *Tale of Igor*, an epic Old Russian poem. While the date of the poem has been disputed for decades, recent linguistic evidence presented by Zaliznjak (2007) has led to the current scholarly consensus that the poem dates back to the late 12th century. In this epic text the lexeme *žirъ* is used three times: once as a noun *žirъ* and twice as an adjective *žirnyj*. The noun is translated by Tvorogov (1997) and Meščerskij (1985) as "abundance, welfare" and the adjective as "abundant, happy".

The noun *žir* in *The Tale of Igor* is translated by Sreznevski as "wealth" (1989 [1893]: I, 2, 875). In these occurrences, the lexeme *žir* is synonymous with the OE lexemes *ead* and *wela*, denoting the meaning of "wealth, abundance, prosperity". To find confirmation for these occurrences, we should turn to the *Dictionary of Russian folk dialects* (Filin & Sorokoletov 1965–2008: IX, 180–188). Here is the selection of meanings that coincide with that of *The Tale of Igor*:

> *žiry* "of good, free, well-fed life" (Dal');
> *žir* "big quantity of something, big number of somebody or something" (Pskov);
> *žir* "living conditions; of good free life" (Arkhangelsk, Olonetsk, Kursk, Tambov, Penza);
> *žirá* "fate, destiny" (Onega);
> *žiréna* "rich person" (Dal');
> *žírnyj* "abundant, high (of water)" (Orlov, Voronež, Saratov, Tambov);
> *žírova, ziróva* "good well-fed life" (Vologda);
> *žírova* "place where the goods are kept" (Olonetsk);

žirovato "(to live) in a big way, beyond one's means" (Ural'sk);

žirovat' "to relax, to feast" (Kolyma, Siberia, Penza); "to produce something in abundance" (e.g. of the earth giving crops; Orel);

žirovat'sia, "to live in prosperity" (Sverdlovsk);

žirovoj "happy, rich" (Kursk, Tambov);

žirovoj "rich person" (Vologda, Onega);

žirov'e "good, prosperous life" (Olonetsk);

žiruxa "prosperous life" (Arkhangel'sk);

žiry "well, freely, prosperously" (Penza, Tambov);

žir'mja "(to live) in prosperity, having everything in abundance" (Arkhangelsk, Novgorod);

žirjak "rich person" (Rjazan').

The notions of "prosperity" and "abundance" should be part of the original meaning of the word *žir*. Miklosich (1886) drew attention to the morphology of *žir*, comparing it with the formation of the word *pir* "feast", which derives from the verb *pit'* "to drink". Thus, *pir* "feast" is a meal where there are drinks in profusion. *Žir* derives from the verb *žit'* "to live". Hence *žir* should be a prosperous life, i.e. a life where food and provisions are abundant. In spite of this analysis, Miklosich merely compared all the personal *žir-* names with the names containing the element *žit'* "*vita*" and *živŭ* "*vivus*" (1927: 62). Subsequent researchers did not question this position.

3.4   Slavic anthroponyms

The dialectal evidence presented in the previous pages indicates that the element *žir-* in Old Slavic personal names did not refer to fat or animal food but rather to prosperity and abundance. In light of this interpretation, the majority of Slavic *žir-* anthroponyms become clear: they have nothing to do with poetic epithets but express parents' wish for profusion and prosperity for their child. Thus the name *Domažir* consists of two themes, *dom* "home" and *žir* "prosperity". The noun *dom* in this name is used in the form *doma*, which corresponds to the defunct common Slavic locative case; so the use of the locative case along with the meaning of the themes allows us to interpret *Domažir* as "prosperity at home" and accept that it is an original formation. An examination of the *zir-* names – except the name *Žiroslav*, which presents a particular phenomenon to be discussed further – reveals that more than a half of them (25 of 47) are monothematic. Although names such as *Žirko* or *Žirjata* may represent the hypocoristic formations of *Žiroslav*, the overall distribution of *žir-* anthroponyms leaves us with the impression that they were originally monothematic.

Thus, the forms *Žiročko, Žirjata, Žirko, Žiroša, Žiroško, Žirox, *Žirna, *Žirneu, *Žirno* are equivalent to the Greek personal names *Plous, Plousia, Plousis, Ploutas, Plouteros, Ploutianos, Ploutias, Ploutina, Ploutinas, Ploutinos, Ploutis, Ploution, Ploutixa, Ploutos* or to the Old Germanic and OE names *Aud, Audo Audic, Edeco, Eada, Eda.*

All Old Slavic compound personal *žir-* names and simple personal *žir-* names with prefixes, except *Žiroslav,* stylistically and semantically express the hope for a well-fed, prosperous life. Examples of such *žir-* anthroponyms are the following: *Dobrožir* "good wealth", *Nažir* "for profusion", *Xot'žir* "eager for prosperity", *Čažir* "hope for prosperity", *Našežir* "our wealth", *Mojžir* "my wealth", *Ranožir* "prosperity at an early time", *Svojžir* "one's own prosperity", *Vratižir* "return prosperity".

Some forms are not immediately recognizable to a modern Slavic speaker. Thus Czech *Kromežir* contains the prototheme *krome,* which in modern Russian means "except". So the interpretation of this name as "except prosperity" seems to go against the common trend (which is to express the hope for a prosperous life). In Old Slavic, however, this adverb corresponded to the noun *kroma* "edge, border" in the locative case (Šanskij 1962–: 8, 404). The meaning of the Old Russian adverb *krome* was "outside, without", so *Kromežir* could be interpreted literally as "wealth beyond the borders", "overwhelming wealth". Names such as *Negožir* and *Žironežko* (diminutive of *Žironeg*) include the common Slavic theme *neg,* which occurs in the Slavic verb *negovati* "to take care of a child". So these names can describe a child as "being abundantly cared for" (Miklosich 1927: 82; Svoboda 1964: 82). Anthroponyms such as *Nežir* "unhappy, not prosperous" and *Nedažir* "do not give wealth" can be interpreted as Slavic protective names, as are *Nekras* "unpretty", *Nenaš* "not ours", *Nemoj* "not mine", *Neljub* "unloved", and so on. Such names were used in families with high infant mortality; their purpose was to deceive the demons, as they suggested that the child was not an attractive prey (Svoboda 1964: 44–45).

Attested occurrences of Slavic names show that the element *žir-* is encountered almost exclusively in the names of non-heroic characters. Though only a few bearers of *žir-* names are identified as commoners on Russian territory and two as serfs in Carinthia, the character of *žir-* names clearly suggests use by the lower strata of society.

It is only the widespread appearance of the name *Žiroslav* in Old Russian sources that conflicts with this statement. The theme *slav-* "fame, glory" is a favorite, and by far the most frequently used, Slavic anthroponomical theme in dithematic names. Originally this theme was used exclusively by Slavic nobility. In Kievan Russia, the Scandinavian dynasty of Rjurik adopted this practice from the Slavic nobility and used the element *slav-* in almost all the Slavic names of their heirs. Taking into consideration this peculiarity of the element *slav-,* and bearing in mind that the *žir-* element is generally used in commoners' names, it becomes clear that

the name *Žiroslav* was a newly coined anthroponym. First of all, it is recorded only in Russia; secondly, it is not present among the Russian Rjurikid princely family names. Before the end of the 14th century, 21 *Žiroslavs* can be probably identified. They are exclusively of noble, though not princely, origin and belong to the retinue of Kievan princes or the Novgorodian elite. Half of these 21 persons are identified with Novgorod. Everything indicates that *Žiroslav* was coined in the Novgorod region in an attempt to ennoble names of the *Žirox*/*Žirjata* type. Here we encounter the same mechanism as described by Bechtel (1917: 372) in the case of the name *Ploutakandros*. Bechtel notes that the initial anthroponym had to be *Ploutax* and that the form *Ploutakandros* results from adding the element of the father's name, in this case *-andros*.

The peculiarity of the form *Žiroslav* is also accentuated by the absence of other compound anthroponyms of the noble type, such as *Žiromir or *Žirovoj, where *mir* "world" and *voj* "warrior" are themes indicating the high origin of its bearer.

The Slavic and Greek use of the theme referring to wealth allows us to apply the same scheme to the Germanic area. Thus, the Germanic element *ead* should primarily form an uncompounded name intended for the lower classes.[14] In Germanic as well as in ancient Greek and Slavic traditions, this name-element possessed no heroic coloring, but it would correspond to a farm family's highest expectations for their child.

4.   Dithematic Germanic anthroponyms containing the *ead-* element

Following this conclusion, Germanic anthroponyms such as *Odewalo* (attested in the 9th century) and *Wallod* (attested in the 7th century) were later creations, inspired by the poetic compounds OE *eadwela*, OS *ôd-welo*, and OHG *ôt-uuála* analyzed in this paper. Before these forms came into existence, simple Germanic anthroponyms like *Auda*, *Eada*, and *Eadic* were transformed in the same way as the

---

14.   This assumption concerns only *ead-* personal names. I do not claim that the Germanic peoples and the Anglo-Saxons used exclusively monothematic names for the lower classes and dithematic for the upper classes (on the genesis of the Germanic dithematic names, see Khallieva Boiché 2015: 116–120). The idea is that several centuries before the Migration Period, the majority of Germanic personal names were monothematic. Dithematic names began to proliferate when the Germanic tribes, expanding from the Baltic Sea southward at the expense of the Celtic-speaking peoples (Mallory & Adams 1997: 218, 344), came into contact with the the Roman Empire (1st century BCE–1st century CE). At this period, the Germans borrowed the Celtic title *rīk* "king, chieftain". It is probable that the necessity to distinguish the new military elite encouraged the Germanic peoples to use the themes *rīk-* and *mar-/mēr-* in the same manner as the Celts did, both as name-elements and as a mark of belonging to the group of rulers.

Slavic name *Žirox* and the Greek name *Plutax*; that is, they were extended by the addition of a second theme in order to ennoble the *aud-/ead-* element.

Leaving aside Odvulf, a legendary ancestor of Theoderic the Great, the first dithematic *aud-* name appears in the middle of the 4th century in the personal name *Audigild* (Reichert 1987: I, 95). The second element of this name is cognate with German *gelten* "to be worth" (Förstemann 1900: 638). It is suggested here that this theme forms a semantic unity with *aud-* and can be understood as "worthy of wealth". Consequently, the name *Audigild* can be compared to those Slavic dithematic anthroponyms in which the first element reinforces the meaning of the second, *-žir*: *Dobrožir* "good wealth", *Xot'žir* "wish prosperity", *Čažir* "expect prosperity". Semantically these names do not convey any heroic values but hope for well-being.

Table 2 offers a list of the earliest attestations of the Germanic *Aud-* personal names extracted from the *Lexikon der Altgermanischen Namen* (Reichert 1987: I, 94, 95, 529–533).

Table 2. A list of the earliest attestations of the Germanic *Aud-* personal names

Name	Century	Gloss
Aud 1	3rd century	"wealth"
Aud 2	3rd century	"wealth"
Aud 3	3rd century	"wealth"
Audigild	4th century	"worthy of wealth"
Odothe	4th century	"servant of riches"
Odovacar	5th century	"warden of riches"
Odovacri	5th century	"warden of riches"
Audefled[*]	Second half of the 5th century	"wealth" + "beauty, brightness" (Böhler 1931: 157)

* *Audofled* is a secondary formation; on the genesis of Germanic feminine anthroponyms, see Khallieva Boiché (2015: 208).

The first dithematic *aud-* name that matches heroic values is *Odovacar* (*Audovacar*; Reichert 1987: I, 531–533). This name belonged to two individuals in the 5th century, one of whom is the famous king of Italy and the second a Saxon chief. Semantically the form *Odovacar* is close to the Greek *Plutarxos* "the lord of the wealth". *Plutarxos*, in turn, would probably be coined in the same way as Slavic *Žiroslav*, an attempt to render the simple name *Ploutos* noble.

Audovacar corresponds to the OE compound *Eadwacer*, translated by Bosworth & Toller as "watchman of property" and occurring in the OE poem *Wulf and Eadwacer*. It is generally held that *Eadwacer* is used in this poem as a personal name. The collocation in itself can be understood as a kenning for a lord. However, no such poetic epithet is conserved in either OE or OS poetry. Germanic poetry

reflects the relationship between the lord and his people, the lord and his band of warriors, or the lord and his realm, whereby the lord is often designated as a protector, but a protector of his people or his territory, not of treasure, wealth, or property. This relationship is reflected in the following OE poetic epithets: *freawine folca* "the friend of the people" (*Beowulf* 430), *weard Scildinga* "the guardian of the Scyldings" (*Beowulf* 229), *helm Scyldinga* "the protector of the Scyldings" (*Beowulf* 371, 456, 1321, 2381), *mundbora minum magoþenum* "the guardian of my young thanes" (*Beowulf* 1480), *rican þeodne* "to the chief of the nation" (*Judith* 11). With the advent of Christianity, these kinds of epithets were increasingly applied to God: thus approximately a fifth of the poetic occurrences of *weard* are encountered in the collocation *heofonrices weard*, literally "the guardian of the heavenly kingdom" (*CASPR*: 1396–1397).

Besides *vacar, Germanic name-elements corresponding to the designation of the lord as ruler, protector, and guardian are: *-mund* (OE *mund* "protection"), *-helm* (OE *helm* "protector"), *-ric* (OE *rice* "powerful", *rica* "ruler"), *-wald* (OE *walda* "ruler"), *-war* (OE *waru* "shelter, guard"; OHG *waran* "to guard, to preserve"; Schramm 1957: 167).

The name-elements *mund-*, *ric-*, *wald-*, and *war-* form the most ancient dithematic personal names, among which we find personal names corresponding to the poetic epithets describing the lord as the protector of his people or his kingdom (Reichert 1987: I; Khallieva Boiché 2015: Annexe 1):

Hunimund (1st century): *hun + mund* "protector of the Huns";
Hariomund (4th century): *harja + mund* "protector of the people";
Reccimund (5th century): *ric + mund* "protector of the kingdom"
    (Böhler 1931: 109);
Laudaric (5th century): *leud + ric* "leader of the people";
Hunirix (5th century): *hun + ric* "captain of the Huns";
Deudorix (1st century) and Theoderic (ve siecle): *þeud + ric* "leader
    of the people";
Chariovald (1st century): *harja + vald* "ruler of the people";
Gumoari (4th century): *gum + war* "guardian of men";
Landawarij (5th century): *land + war* "guardian of land".

Anthroponyms with *helm-* and *ward-* are absent in Reicherts' (1987) list of ancient Germanic names. According to Förstemann (1900: 808, 1538, 1539), they appeared in the 6th century, but the majority of the occurrences listed by Förstermann are from the 8th century (*Liuthelm*, *Lanthelm*, *Landelhelm*, *Thiothelm*, *Deutselm*, *Hariward*, *Dedoard*).

The name *Odovacar* may have been created because it permitted the ennobling of the simple and common name of the king's father Edico by reinterpreting the

existing poetic epithet referring to the lord as the ruler of the people, in that the epithet "the guardian of the people" would be changed into "the guardian of the wealth". This form was followed by synonymous forms such as *Audemund* (7th century), *Audovald* (6th century), *Oderic* (6th century; Reichert 1987: 94–96; 528–533). The anthroponyms *Audoard* and *Authelm* appear only in the 8th century (Förstemann 1900: 808, 1538).

In the OE poetic corpus (*CASPR*, s.v.), the following collocations denoting "the guardian/protector of wealth" could be attested:

hordweard (*Beowulf*, 1047, 1852, 2293, 2302, 2554, 2593; *Exodus* 35, 512;
    Daniel 65);
beaga weard (*Genesis* 2783);
beahhorda weard (*Beowulf* 921);
goldweard (*Beowulf* 3081);
maðma mundbora (*Beowulf* 2779).

The compound *hordweard* "guardian of the treasure-hoard" is used six times in *Beowulf*. On four occasions, this compound refers to the dragon. On two occasions in *Beowulf* (1047, 1852), and in *Exodus* and *Daniel*, it is used as a positive epithet applied to a lord. The same positive use is found in the collocations *beahhorda weard* and *beaga weard* (said of Abraham). However, *goldweard* and *maðma mundbora* again refer to the dragon. It seems that those epithets referring to the dragon are the genuine ones.

An interesting development is observed with respect to these collocations. It appears that the Anglo-Saxons introduced at least two new themes referring to wealth, treasure, and riches into their onomasticon:[15] *beag-* and *hring-*. Moreover, there are Old English personal names that correspond to *beaga weard* epithets: *Beaghelm*, *Beagmund*, *Hringweard*. However, these name-elements were unknown to other Germanic peoples. Furthermore, presumably the Anglo-Saxons started to use both elements only from the second half of the 8th century (two occurrences of *Beagnoth*, one occurrence of *Hringwine*). Other occurrences of *Beag-* names are from the 9th century, and those involving *Hring-* are much later (10th, 11th, and 12th centuries; *PASE*; *DLV*). This confirms that the Anglo-Saxon onomastic system was productive and continued to coin dithematic names that reinterpreted poetic ephithets, as well as mechanically combining anthroponomical themes.

To return to the use of the name-element *ead-*, the argument is that *ead* was such a widely used poetic word that if the compounds *eadweard/ *eadhelm/*eadmund

---

15. The third theme is *streon-* (OE *streon* "treasure"). Two occurrences, *Streonbeorht* and *Streonwulf*, are found only in *DLV* (Insley et al. 2007: 149).

had existed, they would have been conserved in the poetic corpus. All other Germanic names containing the *ead-* element are secondary formations (see Section 1) and are the results of variation of the personal names within the same family. So, in the period before the 9th century the most frequent Anglo-Saxon *ead-* names are *Eadwulf* and *Eadberht* followed by *Eadwald*, *Eadbald*, *Eadwuini*, and *Eadred*.

5.   Conclusion

The close analysis of Old Germanic anthroponomy reveals a name-element that ought originally to have been used in the lower strata of society. This name-element, *auda-/OE *ead-*, denoted "wealth" and reflected parents' hopes that their children would have a prosperous life. Initially this element was used to form monothematic names. The oldest dithematic name *Audigild* "worthy of riches" expressed the same idea as monothematic *aud-* names. Starting from the 5th century, the period marked by the emergence of new Germanic polities, *aud-* enters the stock of anthroponomical themes intended to coin dithematic heroic personal names. This process is seen in the name of *Odovacar*, which is likely to be the first attempt to adapt this commoners' element *aud-* to the heroic values of the Heroic Age. *Odovacar*, followed in the course of time by synonymous formations with the deuterothemes *-wald*, *-ric*, *-mund*, *-war*, *-helm*, and *-ward*, was not a fortuitous combination of two themes. This combination was dictated by the poetic language and the necessity to name the war chief with a praise epithet.

These conclusions find support in a close examination of Slavic and Greek onomastic material, which revealed a common pattern enabling comparison with the Germanic data. The Greek anthroponomical element *plouto-* in personal names refers to wealth and is mostly found in uncompounded names. The dithematic name *Ploutarxos* stands for "the chief of the wealth" and corresponds to the Germanic anthroponyms *Audvald*, *Edric*, *Edmund*, *Edward*, and *Odovacar*, which reinterpret the poetic epithet defining the lord as a ward of his people and his land.

A Slavic anthroponomical equivalent is the theme *žir-*. A survey of this name-element and lexeme led to the conclusion that both originally referred to wealth and prosperity. The majority of personal names formed with the the *žir-* element are monothematic, whereas the less numerous dithematic names of the *Domažir* or *Xot'žir* type, where the first element reinforces the meaning of *-žir*, keep their prosaic non-heroic connotation. The Old Russian anthroponym *Žiroslav*, comprising deuterotheme *slav-* "glory", does not match this pattern. Like Germanic *Odovacar* or *Audvald* and Greek *Ploutarxos*, this name was coined for the new elite.

Data sources

ASPR – Krapp, George & Elliott Dobbie. 1931–1953. *The Anglo-Saxon poetic records: A collective edition*. New York: Columbia University Press.

BT – Bosworth, Joseph & T. Northcote Toller. 1898 [Supplement 1921]. *An Anglo-Saxon dictionary* (revised by Alistair Campbell 1991, suppl. 1995). Oxford: Oxford University Press.

CASPR – Bessinger, Jess & Philip Smith (eds.). 1978. *A concordance to the Anglo-Saxon poetic records*. Ithaca/London: Cornell University Press.

CDRCDS – Kostrenčić, Marko (ed.). 1967. *Codex diplomaticus regni Croatiae, Dalmatiae et Slavoniae*, vol. 1. Zagreb: JAZU.

DLV – Rollason, David & Lynda Rollason (eds.). 2007. *The Durham Liber Vitae*, 3 vols. London: The British Library.

PASE – *Prosopography of Anglo-Saxon England*. 2010. http://www.pase.ac.uk (accessed April, 2014.)

S – Keynes, Simon, Susan Kelly, Sean Miller, Rebecca Rushforth, Emma Connolly, Rory Naismith, David Pelteret, Levi Roach, David Woodman (eds.). 1999–. *The Electronic Sawyer, a revised, updated, and expanded version of Peter Sawyer's Anglo-Saxon charters: An annotated list and bibliography, published by the Royal Historical Society in 1968*. http://www.esawyer.org.uk/about/index.html (accessed April, 2014.)

References

Adams Bellow, Henry. 1936. *The Poetic Edda* (Old Norse–English transl.). Princeton: Princeton University Press. http://www.sacred-texts.com/neu//poe/index.htm (accessed January 2017).

Bechtel, Friedrich. 1917. *Die Historischen Personennamen des Griechischen bis zur Kaiserzeit*. Halle: Niemeyer.

Böhler, Maria. 1931. *Die altenglischen Frauennamen*. Berlin: Dr. Emil Ebering.

Bradley, Sid A. J. (ed. & trans.). 1982. *Anglo-Saxon poetry*. London/Dent: Everyman.

Chantraine, Pierre. 1968–1977. *Dictionnaire étymologique de la langue grecque: Histoire des mots*, 4 vols. Paris: Klincksieck.

Colman, Fran. 1992. *Money talks: Reconstructing Old English*. Berlin: Mouton de Gruyter.

Colman, Fran. 2014. *The grammar of names in Anglo-Saxon England: The linguistics and culture of the Old English onomasticon*. Oxford: Oxford University Press.
doi:10.1093/acprof:oso/9780198701675.001.0001

de Vries, Jan. 1956 [1935–1937]. *Altgermanische Religionsgeschichte*, 2 vols. Berlin: Walter de Gruyter.

Ellis, Hilda. 1943. *The road to Hel*. Cambridge: Cambridge University Press.

Filin, Fedot & Fedor P. Sorokoletov (eds.). 1965–2008. *Slovar' russkix narodnyx govorov* [Dictionary of Russian folk dialects]. Moscow: Nauka.

Förstemann, Ernst. 1900. *Altdeutsches Namenbuch, Personennamen*. Bonn: P. Hanstein's Verlag.

Fraser, Peter & Elaine Matthews (eds.). 1987–2010. *A lexicon of Greek personal names*, 4 vols. Oxford: Clarendon Press.

Gering, Hugo. 1903. *Vollständiges Wörterbuch zu den Liedern der Edda*. Halle: Verlag der Buchhandlung des Waisenhauses.

Gneuss, Helmut. 1991. The Old English language. In Malcolm Godden & Michael Lapidge (eds.), *The Cambridge companion to Old English literature*, 23–54. Cambridge: Cambridge University Press. doi:10.1017/CCOL0521374383.002

Greule, Albrecht. 1996. Morphologie und Wortbildung der Vornamen: Germanisch. In Ernst Eichler, Gerold Hilty, Heinrich Löffler, Hugo Steger & Ladislav Zgusta (eds.), *Namenforschung/Name studies/Les noms propres: Ein internationales Handbuch zur Onomastik/An international handbook of onomastics/Manuel international d'onomastique*, 1182–1187. Berlin: Mouton de Gruyter.

Insley, John. 2007. The Old English and Scandinavian personal names of the Liber Vitae to 1200. In David Rollason & Lynda Rollason (eds.), *The Durham Liber Vitae*, vol. 2, 9–12. London: The British Library.

Insley, John & David Rollason. 2007. English monothematic names. In David Rollason & Lynda Rollason (eds.), *The Durham Liber Vitae*, vol. 2, 165–187. London: The British Library.

Insley, John, David Rollason & Peter McClure. 2007. English dithematic names. In David Rollason & Lynda Rollason (eds.), *The Durham Liber Vitae*, vol. 2, 81–162. London: The British Library

Jakobson, Roman. 1950. Slavic mythology. In Maria Leach (ed.), *Funk and Wagnalls Standard dicitonary of folklore, mythology and legend*, vol. 2, 1025–1028. New York: Funk & Wagnalls.

Jónsson, Guðni (ed.). 1949. *Eddukvæði Sæmundar-Edda* [Poetic Edda]. Reykjavík: Íslendingasagnaútgáfan, HEIMSKRINGLA. http://heimskringla.no/wiki/V%C3%B6lundarkvi%C3%B0a (accessed January, 2017.)

Kelle, Johann. 1889. *Untersuchungen zur Überlieferung, Übersetzung, Grammatik der Psalmen Notkers*. Berlin: Weidmann.

Khallieva Boiché, Olga. 2015. *Imja et name: Aux sources de l'anthroponymie germanique, anglo-saxonne et slave*. Paris: Presses de l'Université Paris-Sorbonne.

Kroonen, Guus. 2013. *Etymological dictionary of Proto-Germanic*. Leiden: Brill.

Lecouteux, Claude. 1997. *Les nains et les elfes*. Paris: Imago.

Lehmann, Winnfried. 1986. *A Gothic etymological dictionary*. Leiden: Brill.

Lühr, Rosemarie (ed.) 2017. *Etymologisches Wörterbuch des Althochdeutschen*, 6 vols. Göttingen: Vandenhoeck & Ruprecht.

Mallory, James & Douglas Q. Adams (eds.). 1997. *Encyclopedia of Indo-European culture*. London: Fitzroy Dearborn Publishers.

Masson, Olivier, Catherine Dobias & Laurent Dubois (eds.). 1990. *Onomastica graeca selecta*, vol. I. Nanterre: Université de Paris X.

Meščerskij, Nikita (ed. & trans.). 1985. *Slovo o polku Igoreve* [The tale of Igor]. Leningrad: Sovetskij pisatel'.

Miklosich, Franz. 1886. *Etymologisches Wörterbuch der slawischen Sprachen*. Wien: W. Braumüller.

Miklosich, Franz. 1927. *Die Bildung der slawischen Personen- und Ortsnamen*. Heidelberg: C. Winter.

Motz, Lotte. 1986. New thoughts on *Völundarkviða. Saga-Book of the Viking Society* 22. 50–67. London: University College London.

Nedoma, Robert. 1988. *Die bildlichen und schriftlichen Denkmäler der Wielandsage*. Göppingen: Kümmerle.

Nilsson, Martin. 1941. *Geschichte der griechischen Religion*. München: Beck.

Ochs, Ernst. 1921. Die Heiligen und die Seligen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 45. 102–112.

Okasha, Elisabeth. 2011. *Women's names in Old English*. Surrey: Ashgate.

Piper, Paul (ed.). 1883. *Die Schriften Notkers und seiner Schule, Psalmen und katechetische Denkmäler nach der St. Galler Handschriftengruppe*. Freiburg: Mohr.

Pokorny, Julius. 1959–1969. *Indogermanisches Etymologisches Wörtbuch*, 2 vols. Bern/München: Francke Verlag.

Redin, Mats. 1919. *Studies on uncompounded personal names in Old English*. Uppsala.

Reichert, Hermann. 1987. *Lexikon der altgermanischen Namen*, 2 vols. Wien: Österreichische Akademie der Wissenschaften.

Rosenfeld, Hellmut. 1969. Der Name Wieland. *Beiträge zur Namenforschung* 4. 53–62.

Šanskij, Nikolaj (ed.). 1962–. Etimologičeskij slovar' russkogo jazyka [Etymological dictionary of Russian language], 10 vols. Moscow: Moscow State University.

Schlimpert, Gerhard. 1978. *Slawische Personennamen in Mittelalterlichen Quellen zur deutschen Geschichte*. Berlin: Akademie Verlag.

Schmid, Hans-Jörg. 2016. *English morphology and word-formation*. Berlin: Erich Schmidt Verlag.

Schmitt, Rüdiger. 1973. *Indogermanische Dichtersprache und Namengebung*. Innsbruck: Institut für Sprachwissenschaf der Universität Innsbruck.

Schramm, Gottfried. 1957. *Namenschatz und Dichtersprache, Studien zu den Zweigliedrigen personennamen der Germanen*. Göttingen: Vandenhoeck & Ruprecht.

Sehrt, Edward. 1962. *Notker-Glossar: Ein Althochdeutsch-Lateinisch-Neuhochdeutsches Wörterbuch zu Notkers des Deutschen Schriften*. Tübingen: Max Niemeyer.

Sehrt, Edward. 1966. *Vollständiges Wörterbuch zum Heliand und zur altsächsischen Genesis*. Göttingen: Vandenhoeck & Ruprecht.

Sehrt, Edward & Wolfram Legner. 1955. *Notker-Wortschatz*. Halle: VEB Max Niemeyer.

Sievers, Edward (ed.). 1872. *Tatian, Lateinisch und altdeutsch*. Paderborn: Ferdinand Schöningh.

Sreznevskij, Izmail. 1893 [1989]. *Slovar' drevnerusskogo jazyka* [Dictionary of Old Russian]. Moskva: Kniga.

Svoboda, Jan. 1964. *Staročeská osobní jména a naše příjmení* [Old Czech personal names and our surnames]. Prague: Nakladatelství Československé akademie věd.

Taylor, Paul. 1963. The structure of Völundarkviða. *Neophilologus* 47. 228–236.
doi:10.1007/BF01514967

Tvorogov, Oleg (ed. & trans.). 1997. *Slovo o polku Igoreve* [The tale of Igor]. St-Peterburg: Nauka.

Vajda, Edward. 2001. Semantics. In Edward Vajda, *Introduction to linguistics, fall quarter 2001 course materials*. Bellingham: Western Washington University. http://pandora.cii.wwu.edu/vajda/ling201/test3materials/semantics.htm (13 November, 2015.)

Weisweiler, Josef. 1923. Beiträge zur Bedeutungsentwicklung germanischer Wörter für sittliche Begriffe. *Indogermanische Forschungen* 41. 304–368.

Woolf, Henry Bosley. 1939. *The Old Germanic principles of name-giving*. Baltimore: The Johns Hopkins Press.

Zaliznjak, Andrej. 2007. *Slovo o polku Igoreve. Vzgljad linguista*. [The tale of Igor: A linguist's point of view]. Moscow: Rukopisnye pamjatniki Drevnej Rusi.

News and *relations*

Highlighted textual labels in the titles of early modern news pamphlets

Carla Suhr
University of Helsinki

This paper is a pragma-philological examination of the textual labels and visual features in the titles and title pages of 53 early modern sensationalist news pamphlets. The analysis shows that certain textual labels are favored over others, that textual labels were not highlighted until the 1640s, and that before that time the overall visual layout of the title page was a more important genre marker than textual labels. These developments in the conventionalization of textual labels can be tied in with increasing literacy and access to printed texts of the unlearned masses that were the primary readership of these pamphlets.

Keywords: historical linguistics, pragmatics, pragma-philology, English, news pamphlets, genre, textual labels, titles, illustrations

1. Introduction

The book historian Eleanor Shevlin claims that "titles embody the potential to illuminate not just individual works, but reading processes, authorial composition, publishing practices, marketing trends, and generic transformations as well" (1999: 43–44). In the past few years, linguists – especially in the field of pragmatics – have picked up on several of these aspects of titles. They have taken notice especially of the potential of textual labels given in titles to carry genre expectations (see Bach 1997; Suhr 2011; Ratia 2013). At the same time, linguists have also begun to recognize that visual features of a page can carry linguistic meanings or aid the reading comprehension of a text (Bach 1997; Suhr 2011; Carroll et al. 2013; Salmi 2013; see also the contributions in Meurman-Solin & Tyrkkö 2013). Other studies have acknowledged the competition between marketing trends and what Shevlin (1999: 43) calls the contractual nature of the title (see Tyrkkö et al. 2013).

The importance of the textual labels of early English news publications was hinted at already in 1996 by the historian Joad Raymond in his study of newsbooks, which were short, quarto format periodical news publications of the 1640s. He cursorily notes that the term 'news' had acquired a potentially pejorative connotation from its use in earlier occasional sensationalist news pamphlets focusing on anomalous phenomena such as monstrous births, witchcraft, trials and executions, or natural disasters (1996: 158–159; see Raymond 2003: 98–128 for a description of sensationalist news pamphlets). Because of this taint, Raymond argues, *news* was more likely to be used in titles for foreign news, whereas 'proper' domestic news – which began to be printed only in the 1640s – would appear under the label of *intelligence* or *information*, which implied sophistication and seriousness but also expertise of the editor in sifting through meaningless news and interpreting it to inform the judgment of the readers (1996: 158–163). However, there has been no systematic study of the textual labels of sensationalist news pamphlets from before 1640 that could corroborate Raymond's claim that *news* was such a common label for these texts that its use could not be considered for other types of domestic news reports.

My study of the visual and linguistic features of the title pages in a corpus of English witchcraft pamphlets (Suhr 2011: 125–142), one type of sensationalist news pamphlet, suggests that it was precisely around the 1640s when self-descriptive textual labels such as *news* or *relation* started to carry content and genre expectations. Earlier pamphlets did not highlight textual labels on the title page because visual features such as illustrations and layout marked a text as popular, or as a text aimed at the growing semi-literate middle and lower classes that benefited from increasing basic literacy. The formative years from the late 16th century up to the 1640s, the study argues, were a period when the new genre became conventionalized and reading became a common enough skill that readers could use highlighted textual labels to guide their expectations of the content and structure of a text rather than visual features of the title page (Suhr 2011).

This study builds upon these earlier studies. On the one hand, it tests Raymond's claim about the 'contamination' of the textual label *news* because of its prevalence in the titles of sensationalist news pamphlets before the 1640s. This is be done by charting the repertoire of textual labels and by identifying any diachronic patterns of over 50 sensationalist news pamphlets from the period 1580–1700 that deal with monsters, storms, and encounters with the devil. At the same time, this study tests whether the turn from visual to verbal primacy observed in witchcraft pamphlets and dated to around the 1640s can be seen in the other sensationalist news pamphlets; to that effect, it records the presence or absence of illustrations and visual highlighting of linguistic items. Both the conventionalization of textual labels and how they were marked in popular news pamphlets can be tied to the

larger cultural issues of literacy, reading habits, and changing print culture. This pragma-philological approach thus not only sheds light on the early development of genre conventions in news writing but also demonstrates how the pragmatic study of texts can be advanced by taking account of visual aspects of the textual page in the analysis of linguistic features that help readers structure and process texts (for syntactically oriented approaches to genre analysis, see the contributions by Moessner, Gather, and Nakayasu in this volume).

This paper is organized as follows. The following section describes the main findings of previous studies regarding textual labels in Early Modern English pamphlets. Section 3 introduces the materials of the current study and outlines the method used to analyze textual labels and their relevance for readers, while Section 4 presents the results of the study. Whether the claims of Raymond (1996: 158–163) and Suhr (2011: 125–142) are supported by the results of the current study is discussed in Section 5. Section 6, then, concludes with remarks about the broader implications of this study and suggestions for further research.

2. Previous linguistic observations on textual labels in Early Modern English pamphlets

There are three linguistic studies that consider textual labels in Early Modern English pamphlets.[1] Two of them (Bach 1997; Claridge 2000) focus on the period starting in 1640, so they cannot shed light on Raymond's claim about the earlier use of *news* as a textual label for sensationalist news pamphlets. In addition, the majority of the pamphlets these studies investigate are not sensationalist news pamphlets but polemical pamphlets that argued for particular political, religious, or scientific viewpoints rather than reported on news. These two studies are therefore of limited use in terms of determining the variety of textual labels for specifically sensationalist news pamphlets from 1580 onwards, or in testing the 1640s watershed hypothesis. They do, however, provide interesting points of comparison for the present study, so I summarize their main suggestions and findings regarding

---

1. After the submission of this chapter, two articles appeared that include an analysis of the metadiscursive use of what I would call textual labels in historical English news discourse. Brownlees (2015: 5) describes the terms he analyzes as nominal forms of metadiscourse terms; his analysis focuses on such terms used in the title pages and in editorial addresses and comments of periodical news publications of the 17th century, which are closely related to sensationalist news pamphlets (called 'occasional news pamphlets' by Brownlees 2015: 13). Bös (2015) investigates the quantitative development and potential semantic and functional changes of 25 news-related key terms, not all of which are textual labels. Her material is comprised of news stories from the period 1700–2000.

textual labels before describing the study of textual labels of witchcraft pamphlets (Suhr 2011: 125–142) that this study replicates.

Bach (1997) contains an in-depth analysis of the textual labels – he uses the term 'text-type label' – of almost 500 pamphlets included in the Thomason tracts, dating from the period 1640–1661.[2] The study analyzes the kinds of textual labels found in these texts and shows how the clustering of labels signals participation in chains of pamphlets and controversies dealing with specific people, groups of people, or institutions. Furthermore, Bach shows how textual labels in the titles, supported by typographical highlighting and sometimes illustrations, can be associated with certain schemas (LIST, SERMON, LIFE-AND-DEATH, and TRIAL) that helped readers (and listeners) to keep track and make sense of the actual pamphlet text; textual labels, then, activate or trigger in the reader certain expectations of the contents and structure of the text to follow. Bach also identifies certain textual labels (*inventory*, *funeral*, *fair*) that worked as *imagines et loci* mnemonic devices that could be used to help remember the pamphlet text. *News* is not a label that Bach discusses either as a schema or a mnemonic device. Bach's findings about schemas or mnemonic devices are not directly applicable to sensational news pamphlets, because many of the textual labels in his material simply do not appear in sensational news pamphlets. In fact, Bach does not provide a complete inventory or statistics of the textual labels found in his material, so there are no direct points of comparison between his discussion of textual labels and the results of this study. However, the importance of textual labels, emphasized as it is by typographical highlighting, certainly seems to support Raymond's point about avoiding the term 'news' as inappropriate for the newsbook genre (1996: 158–159) and also my own (2011) claim about a shift in the 1640s towards verbal genre marking in witchcraft pamphlets.

Claridge (2000) mentions what I term textual labels in her study of interactive features in pamphlets included in the *Lampeter corpus of Early Modern English tracts* (2003), which covers the period 1640–1740. According to her, textual labels point to established genres or text-types or speech acts by describing "what kind of text the pamphlet is or what function it has" (2000: 28), but they also imply who the intended addressee or reader of the text is because they are expressions that require a target, be it a person or another pamphlet. Claridge links this characteristic with a wish to make a more personal connection with the reader, visible especially in the frequency of the textual label *letter* (which has the added benefit of increased authenticity and credibility when the addressee is an important person) and in texts

---

2.   The Thomason tracts are a collection of pamphlets, books, and newspapers printed between 1640 and 1661. Originally, the texts were collected by the London bookseller George Thomason; they are now held in the British Library (see http://www.bl.uk/reshelp/findhelprestype/news/thomasontracts/ (9 October, 2017)).

written as dialogues (2000: 28–29). This is an intriguing hypothesis, and it is worth testing whether the textual labels in sensationalist news pamphlets show a similar concern for personal connection. A comprehensive inventory of textual labels in her material was beyond the scope of Claridge's study, but in fact a list of these labels is available in the manual of the *Lampeter corpus*, and it is made use of in Section 4.

Claridge (2000) and Bach (1997) do not differentiate between what I call polemical pamphlets and sensationalist or popular news pamphlets. However, this is an important distinction to keep in mind, for the intended audiences of the two kinds of texts were very different: to produce and to follow the sophisticated argumentation full of learned references found in polemical pamphlets, which often evolved into controversies addressing previous texts on the same topic, required a level of literacy beyond the means of the majority of the population (see, e.g., Brayman Hackel 2005: 57–63 and Thomas 1986 on different kinds of literacies). Sensationalist news pamphlets had a much wider audience consisting of increasing numbers of newly semi-literate readers from all but the lowest levels of society (Clarke 1983: 22–23) – and the practice of reading popular texts out loud brought even the illiterate within the potential audience of popular news pamphlets (see, e.g., Fox 2000: 36–39; Suhr 2011: 15–26).

The differences in the level of education of the audiences of polemical pamphlets and sensationalist news pamphlets is very likely to be reflected in their repertoires of textual labels. Learned readers were familiar with a very wide range of genres and the textual labels attached to them through their classical education, whereas the barely literate readers of sensationalist news pamphlets had contact with a much more restricted range of genres that possibly derived from the oral genres they were used to listening to. For such semi-literate readers, building up enough familiarity with reading texts to be able to assign certain kinds of content-related and structural expectations to different textual labels located in the title of a text takes time, especially when texts are still in the process of forming their own textual conventions, as was the case with sensationalist news pamphlets. Neither Claridge (2000) nor Bach (1997) can shed light on the repertoire of textual labels in pamphlets during these formative years, and in fact Suhr (2011: 128–130) only makes general observations about the kind of information included in titles rather than systematically charting textual labels in witchcraft pamphlets from before 1643. The art historian Ruth Samson Luborsky (1987) claims that woodcut illustrations were a generic convention of romances, a popular genre, and Suhr (2011: 123) suggests that this reliance on a visual feature of texts to mark genre also applies to 16th-century witchcraft pamphlets and could well apply to other popular genres. Existing studies of illustrations in all kinds of pamphlets (Davies 1986) and ballads, chapbooks, and penny chapbooks (Watt 1991) up to the year 1640 certainly lend credence to this

supposition, as they make note of the continued importance of illustrations in these texts – though they have not analyzed the precise quantities of illustrations in them.

In fact, I argued earlier (Suhr 2011: 125–142) that the reliance on visual genre marking was retained on the title pages of witchcraft pamphlets until the 1640s, about half a decade longer than in the actual texts. Therefore it very likely did not matter what kind of descriptive textual labels witchcraft pamphlets carried in their titles. The text on the title page was arranged in a visually pleasing manner using different typefaces and decreasing type size, for instance by organizing the text into shapes such as an inverted pyramid; illustrations were expected components of title pages even after they disappeared from the inner pages of pamphlets. No attention was paid to highlighting textual labels that could guide a reader's expectations about the content and style of the pamphlet, with the result that textual labels were harder to identify cognitively for prospective buyers looking at title pages. It seems that, rather than the title or the textual labels embedded in the title, it was the presence of woodcut illustrations and a visually pleasing layout on the title page that were the primary signal for prospective buyers that the text would appeal to a popular audience. According to the study, the 1640s appears to be a watershed: this was the decade when the earlier visual primacy of the title page gave way to the practice of typographically highlighting the words and expressions that label the genre and topic of witchcraft pamphlets. This shift signaled that textual labels now had meaning for both producers and readers of the texts, exemplified also by contemporary newsbook editors' distinctions between the labels *news*, *intelligence*, and *information* that Raymond notes (1996: 158–163).

My earlier study (2011) was limited to one genre of sensationalist news pamphlets only, and the material consisted of only 36 texts. The current study aims to test whether the watershed hypothesis applies also to 53 sensationalist news pamphlets spanning approximately the same time frame and focusing on three other topics: monsters, storms, and encounters with the devil. At the same time, the current study charts the range of textual labels in these texts as well as identifying any emerging trends in labeling. The following section describes the materials and methods of this study in more detail.

3.   Materials and methods

The primary material of this study consists of a small-scale corpus of 53 sensationalist news pamphlets reporting on monsters, including abnormal births (19 texts), storms (15 texts), and encounters with the devil (19 texts) (see Kay & Alexander in this volume for an analysis of the supernatural in English metaphors). They were located in the *Early English Books Online* database using the subject search

keywords 'monsters', 'storm', and 'devil'. These keywords were chosen because they were mentioned as especially frequent topics (alongside witchcraft) of sensationalist news pamphlets in historical overviews of the genre (Clarke 1983; Raymond 2003). Accordingly, the texts included in the corpus can be considered representative of the genre. Note that the initial search yielded far more than 53 texts; the final number was obtained after discarding multiple copies and texts belonging to other genres such as ballads, broadsides, and treatises. Some of the texts are translations of foreign news pamphlets, but the majority of texts deal with singular, 'wonderful and true' events that took place in England.

The selected texts reflect developments in contemporary printing practices. The earliest text is from 1580 and the latest from 1699, so the texts cover roughly the same period as the corpus of 36 witchcraft pamphlets used in my earlier work (2011: 58–60). As with the witchcraft pamphlets, the texts are not spread evenly over the period, reflecting the reality of contemporary publication. For example, Raymond's (2003: 109) observation that news of abnormal births were rare in the 1620s and 1630s but increased again in the 1640s and 1650s correlates with a gap from 1620 to 1640 in the monster pamphlets of the corpus. In general, the majority (40 texts) of the corpus texts are from after 1640: there are only 3 texts from the 16th century, and a further 10 texts from the first four decades of the 17th century. This skewing towards the end of the period is also present in the witchcraft pamphlet corpus, though not quite so clearly (13 of the 36 texts are from before 1640). This is a general feature of Early Modern English publishing, for the number of printed texts, especially short texts, increased dramatically in England from the 1640s onwards. The explosion of print has been connected to the relaxation of censorship with the outbreak of the English Civil War but also with increasing literacy and an increasing demand for news (Raymond 2003: 161–172).

There is one important difference between witchcraft pamphlets and the other sensationalist news pamphlets considered here: almost all the witchcraft pamphlets report on witchcraft trials. As a result, their textual labels are often inherited from parts of trial proceedings regardless of whether the source of information is actual trial documents or eyewitness reports from a trial. This is not the case with the sensationalist news pamphlets dealing with monsters, storms, and devil sightings, which rely on word-of-mouth accounts, eyewitness statements, or sometimes letters as sources. It can therefore be expected that the other sensationalist news pamphlets will have a wider range of textual labels than witchcraft pamphlets, or at the very least that textual labels related to trial proceedings (e.g. *examination*, *confession*) will not be found in the other news pamphlets.

After an initial look at the title pages in the corpus, I divided the texts into two categories: texts that contain typographical highlighting and texts that do not make use of typographical highlighting. Typographical highlighting includes features

such as typeface switches (usually from Roman to italics or vice versa, though black-letter is found as well), larger type size, or the use of all capital letters, all of which have the effect of making words or expressions stand out from the rest of the text. These highlighted items are what readers would see first, and they would therefore very likely play an important part in guiding the readers' expectations of the contents and perhaps also of the structure of the text itself. At least in the 17th century, printers were very aware of the importance of emphasis through typographical highlighting. This is evident in Moxon's guide for compositors, where he writes the following:

> He [the compositor], as aforesaid, judiciously reads his *Title Page*, and considers what *Word* or *Words* have the greatest Emphasis in it. If many *Words* precede the Emphasis, he considers whether it be best to make one or two *Lines*, or more of them, by electing a *Body* bigger or less to *Set* the precedent *Matter* in, and whether any of these *Lines* ought to be *Indented*, either at one end or both, viz. *Set* in the middle of the *Line*. And what Words of Emphasis come in that precedent *Matter*, that he may *Set* them either in *Capitals, Roman, Italick* or *English* [i.e. blackletter]; and at last bring the great Emphasis, which is generally the *Title* or *Name* of the *Book* in a *Line* by it self, and just fill it if he can; which he has some helps to do, by the great *Bodied Letters* of the *Lower Case*, or else by *Capitals, Roman, Italick* or *English*, of a proper *Body*, which best pleases his fancy, or is in present mode.
>
> (Moxon 1683: 221, italics in original)

What is evident from Moxon's instructions to compositors is that typographical highlighting as a way of emphasizing words that were considered most important was normal practice in his time. This was not always the case, as Suhr (2011: 125–142) has shown at least for witchcraft pamphlets; as mentioned in Section 1, one aim of this study is therefore to examine whether the practice of emphasizing important words in the titles of sensationalist news pamphlets became the norm in the 1640s.

In addition to highlighting, I took note of whether the title pages contained illustrations. Both of these visual features of title pages, typographical highlighting and illustrations, are important for determining if other sensationalist news pamphlets show a turn from visual to verbal considerations.

Once the texts were divided into two categories depending on whether their title pages contained highlighted elements, I made an inventory of the various textual labels found in the texts. Using the functional definition of Claridge (2000: 28), I identified textual labels as terms that tell "what kind of text the pamphlet is or what function it has." Claridge's definition derives from the one for the term 'text type and genre self-descriptor' used in the annotations for texts in the *Lampeter corpus*. According to the compilers, these self-descriptors are terms "explicitly given on the title page or somewhere else in the text," and they were only accepted if they were an established genre term (e.g. *sermon*) or a publication type (e.g. *pamphlet*), or if an expression on the title page could be interpreted as a speech act term in a wide

sense (e.g. *answer*) (Claridge 2003: Section 3.2.2). I was, however, more restrictive in my criteria and focused only on textual descriptors found on the title page.

The inventory of labels that I compiled makes a distinction between *primary* and *secondary* textual labels as well as *headlines* and *topical* labels, all of which can be seen in the title page of a 1678 pamphlet reporting on a man possessed by the devil (see Figure 1). It is important to note that all but primary textual labels can occur multiple times on a title page, and in fact combinations of labels are common. Primary textual labels are the first textual labels assigned to a text – in Figure 1 the primary textual label is *news* – and I would argue that it is these textual labels that readers rely on most when determining their expectations of the content and structure of the text. Secondary textual labels are supplementary labels that follow one or more labeling elements – a primary textual label, a headline or another secondary textual label – and they are preceded by tag words such as *or*, *being*, *shewing*, or *with*. These secondary textual labels can be typographically highlighted, but this is rare. Instead, it is strikingly common to highlight the tag words that precede the secondary labels by centering the single word on a line and using all capital letters. Secondary textual labels, then, indicate that additional information is forthcoming; often secondary textual labels are followed by quite detailed descriptions of the contents of the pamphlet. This is also the case in Figure 1: the secondary textual label *relation* comes after the tag *being*, and the narrative of the pamphlet text is summarized after the label.

Headlines are catchy phrases such as "VICTORY over the DEVIL" in Figure 1 or more informative expressions such as "The *KENTISH* MIRACLE" that fit into one or at most two lines. They are intended to catch a potential reader's interest. They usually start the title, but they can also follow another headline or a primary textual label. Headlines are usually typographically highlighted, which makes them stand out from the text. Headlines often replace primary textual labels, though there are a few instances where a headline in small type is followed by a textual label in much larger type. In these cases, I count the textual label as a primary label rather than as a secondary textual label simply because the visually highlighted textual label catches the reader's eye first rather than the headline in smaller type at the top of the page. Finally, title pages may have topical labels that tell something about the topic (e.g. *wonder, tempest*) of the pamphlet or where the described event took place (e.g. *Goswell-Street* in Figure 1). Topical labels always need to be typographically highlighted, i.e. they have to stand out from the surrounding text, in order to be included in the analysis. The difference between headlines and topical labels is that headlines are multi-word expressions whereas topical labels consist of just one word.

Once information about the different kinds of labels found in other sensationalist news pamphlets had been collected, the next step was to compile an inventory of primary and secondary textual labels that could be compared to the textual labels found in witchcraft pamphlets (Suhr 2011) and the pamphlets included in the

Strange and Wonderful

NEWS

FROM

GOSWELL-STREET:

OR, A

VICTORY over the DEVIL.

BEING

A true Relation how a Person, living at the House of *Francis Jordan* at the Sign of the Huntsman and Hounds, near *Mount-Mill* in *Goswell-street*, having for three years last past lain under an Evil-Tongue and lamentable Fits, generally judged to proceed from Witchcraft, and was in a lamentable condition, and her flesh was as if it were tore off her bones.

With the strange Noises heard at that time, and how the Spirit struck the man of the House a grievous blow on the head, which for some Hours occasion'd great pain to him, but now he is recovered.

The Truth of this Relation is and will be attested by *Francis Jordan* at the Sign of the Huntsman and Hounds, near *Mount mill* aforesaid.

Susan Shawe, } { *Rachel Hopkins.*
Margaret Flamstead. } { *Ralph Jordan.*
And several other persons.

With Allowance.

LONDON: Printed for *D. M.* 1678.

Figure 1. Title page of a sensationalist news pamphlet from 1678. The original document is held in the Guildhall Library, City of London. The image is published with permission of Proquest. Further reproduction is prohibited

Lampeter corpus (Claridge 2003: Appendix V). This inventory was used to identify potential long-term trends in textual labeling of sensationalist news pamphlets. The analysis of labels was then refined to focus on elements that are highlighted visually, so that it was possible to see whether the 1640s is a watershed in terms of switching from visual to verbal considerations in title page design, as is the case with witchcraft pamphlets.

4. Results and analysis

4.1 Textual labels

The first aim of this study is to compile an inventory of textual labels found in sensationalist news pamphlets. The 53 pamphlets considered in this study contain 15 unique textual labels and 21 different headlines. Only four of the texts contain neither textual labels nor headlines, but instead make use of highlighted topical labels. Altogether, textual labels are used 72 times as either primary or secondary textual labels. Table 1 provides a breakdown of the textual labels and the number of times (in parentheses) they are used in the corpus texts' title pages.

Table 1. Unique textual labels used in the title pages of sensationalist news pamphlets

relation (32), *news* (13), *account* (9), *discourse* (3), *history* (2), *narrative* (2), *report* (2), *letter* (2), *view* (1), *declaration* (1), *testimony* (1), *examination* (1), *description* (1), *enquiry* (1), *miscellanies* (1)
Total: 15 types, 72 tokens

Clearly, the most common textual labels are *relation*, *news*, and *account*, 75% of all occurrences; the remainder of the textual labels are only used once or twice each. It would seem, then, that though there was a fairly wide range of textual labels that could be used in sensationalist news pamphlets, these three textual labels are preferred over others. Most of the labels could be said to be interactive in the sense that they imply a target (e.g., you *relate* a story to someone; see Claridge 2000: 28–29).

Looking at the distribution of these labels over the categories of primary and secondary textual labels (Table 2), *news* stands out, for it is strikingly rare as a secondary textual label. Conversely, *relation* and *account* were more often used as labels for supplementary information, though their distribution is not as distinctive as that of the label *news*. *Relation* and *account* were perhaps considered more neutral or less informative than *news*, which is why they may have been relegated more easily to secondary textual label status. None of these terms fit the schemas Bach (1997) identifies for polemical pamphlets, but they do fit into van Dijk's narrative

Table 2. Primary and secondary textual labels used in the title pages
of sensationalist news pamphlets

Primary textual labels	*relation* (12), *news* (12), *account* (3), *history* (2), *narrative* (2), *view* (1), *report* (1), *declaration* (1)
	Total: 8 types, 34 tokens
Secondary textual labels	*relation* (20), *account* (6), *discourse* (3), *letter* (2), *report* (1), *testimony* (1), *examination* (1), *description* (1), *enquiry* (1), *news* (1), *miscellanies* (1)
	Total: 11 types, 38 tokens

macrostructure (1980: 116), which can be found in all kinds of genres. I argue that *relation* and *account* were already established labels in other genres (e.g. in travel literature) when sensationalist news pamphlets began to be printed, whereas *news* was a new 'unattached' label. The prominence of *news* as specifically the primary and possibly most visible textual label may be why it becomes associated with sensationalist news pamphlets. Raymond (1996: 1) points out that *What news?* was a question asked so frequently in early modern England that it became the subject of satire; pamphlets labeled *news* answered the question on everyone's lips before it was asked. There is very tentative support for Raymond's (1996: 158–159) suggestion that the association between *news* and sensationalism was already in place by 1640, as 3 of the 13 corpus texts from before that date contain *news* as a primary textual label. However, this is hardly conclusive evidence.

As mentioned before, the textual labels found in witchcraft pamphlets are predominantly associated with trial proceedings, because the texts often reported on trials. Only one such label, i.e. *examination* (*testimony* is not used in the legal sense here), is found in other sensationalist news pamphlets. Of the non-trial-related textual labels used in witchcraft pamphlets, three labels stand out: *relation* is used 14 times, and *news* and *account* are each used 5 times. Just as in other sensationalist news pamphlets, *news* is only used as a primary textual label (with one exception), whereas *relation* and *account* are found both as primary and textual labels. Other textual labels used only once or twice include *treatise, rehearsal, discovery, description, life (and death), answer*, and *history.* With a few exceptions, all of the non-trial-related labels are only used in texts printed in the latter half of the 17th century, when witchcraft pamphlets started to report on suspicions of witchcraft and on unexplained phenomena attributed to witchcraft when witchcraft trials began to become rare. When the frequencies of the textual labels in witchcraft pamphlets (excluding trial-related labels) and other sensationalist news pamphlets are combined, a clear trend emerges in favor of just three textual labels: *relation, news*, and *account*.

On the basis of the *Lampeter corpus*, a number of observations can also be made about textual labels used in polemical pamphlets printed after 1640, though no distinction can be made between primary and secondary textual labels. The manual

of the *Lampeter corpus* provides a list of 50 different 'genre self-descriptions' used altogether 169 times in the 120 texts of the corpus (Claridge 2003: Appendix V). Of the three most frequent labels found in sensationalist news pamphlets (*news*, *relation*, *account*), the label *account* is used 8 times – it is the seventh-most frequent one – in polemical pamphlets, *relation* is only used 3 times, and *news* is only used once. However, a closer look at the titles of the texts in which these labels occur reveals that the texts are in fact sensationalist news pamphlets rather than polemical ones, and adjusting the figures leaves only three instances of *account* and one instance of *relation*. Polemical pamphlets in the *Lampeter corpus*, then, seem to favor other textual labels. Many of the most frequently used textual labels of polemical pamphlets indicate a literate genre (*discourse, treatise*), though *answer, letter*, and *sermon* are not as restricted.

Some other tentative trends emerge in the textual labeling of sensationalist news pamphlets when a diachronic perspective is added. Of the most frequent textual labels, *news* is evenly spread over the 17th century, while *relation* picks up only from 1640 onwards, after which it is used consistently until the end of the century. Intriguingly, *account* appears to have an extremely restricted lifespan, found almost exclusively in texts from the 1680s, but this is most likely just coincidence. Headlines are not evenly spread in title pages, as they cluster into one longer period between 1643 and 1675 and a shorter spate in the 1690s. Finally, from about 1675 onwards, there is a notable increase in topical labels combined with a tendency of clustering in texts. Often topical labels are emphasized as prominently as the textual labels that accompany them, and in a few cases topical labels are emphasized at the expense of the primarily textual label. The increasing importance of topical labels alongside textual labels is so pervasive that it must be motivated by something. I suggest that the reason is competition for buyers: sensationalist news pamphlets (and also the first newspapers) were now so ubiquitous that highlighting a textual label was not enough to garner readers' interest without topical labels to catch their attention.[3]

---

3. The top three textual labels used in sensationalist news pamphlets (*news, relation*, and *account*) were also found in periodical news publications (Brownlees 2015). However, as Brownlees does not provide information about the frequency or the diachronic development of these terms (or the other terms he analyzes), no comparison can be made between sensationalist news pamphlets and periodical news publications. In news stories in the period 1700–2000, *news* is used quite consistently and frequently throughout the period, whereas the usage of *account* drops drastically after 1830. *Relation* was not included in the set of key terms investigated. By far the most frequently used term in news stories was *advice(s)* (Bös 2015), but that term too drops in the mid-19th century. Though the two studies by Brownlees (2015) and Bös (2015) are not directly comparable to the present one, they do indicate that the textual labels *news*, *relation*, and *account* were within the repertoire of labels in historical news discourse in general, possibly all the way up to the mid-19th century.

4.2   Visual highlighting

My second aim is to examine whether sensationalist news pamphlets about monsters, storms, and encounters with the devil exhibit a turn from visual to verbal title pages in the 1640s similar to the one noted for sensationalist witchcraft pamphlets (Suhr 2011: 125–142.) The first step is to calculate how many texts out of a total of 53 highlight textual labels, headlines, and topical labels; how many texts do not make use of typographical means for emphasizing words; and how many texts have illustrations on their title pages. The results are presented in Table 3; recall that texts could have various combinations of these features on their title pages.

Table 3. Numbers of texts with highlighting of textual elements and illustrations

Textual labels highlighted	25 texts
Headlines highlighted	13 texts
Topical labels highlighted	22 texts
No highlighting	11 texts
Illustrations	12 texts

Visually prominent title pages had illustrations on them, and they did not highlight textual labels, as a pleasing visual appearance of the title as a whole was more important than individual words. According to the watershed hypothesis, these kinds of title pages are expected to be dominant in sensationalist news pamphlets until the 1640s, when verbally prominent title pages appear and quickly replace the earlier type of title pages. The new type of title page no longer carries illustrations, and important verbal elements (textual labels, headlines, and topical labels) are highlighted typographically to stand out from the rest of the title. The 53 texts were therefore divided into three groups according to their publication dates: pre-1640 (this period actually ends in 1628, as there are no texts from the 1630s), 1640s, and post-1649. If the watershed theory is to hold, the texts in the pre-1640 group are expected to have illustrations but no highlighting, the 1640s group to mix verbal and visual title pages, and the post-1649 group to lack illustrations but at the same time to contain highlighted textual elements.

The pre-1640 group contains 13 texts. The 1640s group initially contained 8 texts, but the last three texts of the decade fit the expected characteristics of the post-1649 group, so the time frames of these two groups were adjusted to 1640–1645 and post-1645. The middle group thus ended up with only 5 texts, while the post-1645 group contains 35 texts. Figure 2 shows how many texts in the pre-1640 and post-1645 group do *not* behave as expected.
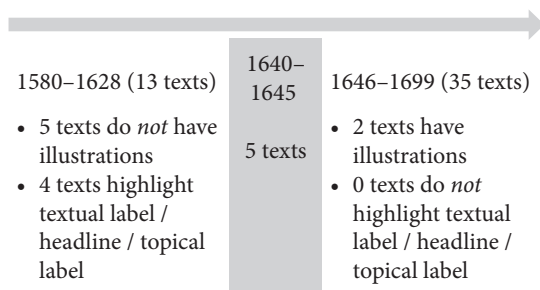
Figure 2. Characteristics of title pages before 1640 and after 1645

The post-1645 group, consisting of two-thirds of all the texts, fits the mold of verbally prominent title pages very well, since all of the texts highlight textual elements, and only two texts (one from 1655 and the other from 1680) have illustrations. The pre-1640 group, which should emphasize overall visuality, seems to fit the expected pattern less neatly. Two of the texts that do not have illustrations do have large ornamental devices instead, so only three texts lack an image of some sort. Furthermore, in three of the four texts with highlighted textual elements, the highlighting may have been unintended, as the textual element appears on a line of its own perhaps simply because it did not fit into the preceding or the following line. The lines of text in visually prominent title pages decreased in size and alternated typeface, so single textual labels on the first line are particularly difficult to interpret. Only one text (from 1613) of the thirteen texts in this group wholly runs counter to expected characteristics, in that it lacks an illustration as well as highlighted textual elements. When we take into account these cases of equivocal categorization, the pre-1640 group fits the expectations rather well, and therefore the watershed hypothesis seems to hold for all sensationalist news pamphlets.

Of the five texts in the middle group, one text (from 1640) fits the pre-1640 group, two texts (from 1643 and 1645) fit the post-1645 group, and two texts (from 1642 and 1645) mix the features of the two groups. Of course, since there are no texts from the 1630s, the middle group could also be extended to start in 1630 (this is also the case with witchcraft pamphlets). However, a study of pamphlet illustrations up to 1640 (Davies 1986) makes no mention of a decrease in pamphlet illustrations in the 1630s, so it seems likely that the visually oriented title pages continued until the 1640s.

5.   Discussion

This study tested two claims regarding textual labels in Early Modern English sensationalist news pamphlets. The first is that *news* was a term associated with sensationalist news pamphlets (Raymond 1996: 158–159), and was therefore avoided in newsbooks. The analysis of textual labels found in 53 sensationalist news pamphlets showed that *news* is indeed the second-most frequently used textual label and that it was the only textual label that was predominantly used as the most prominent primary textual label. *News* was evenly spread through the period investigated here, 1580–1700, though its occurrence before 1640 in only 3 out of 13 texts is at best tentative support for Raymond's claim. The repertoire of textual labels was actually quite small, especially considering the practice of using several labels in one title page: there are only 15 unique textual labels in the 53 pamphlets. Furthermore, three of the labels (*news, relation*, and *account*) are clearly preferred over the others.

However, as the sample size is still quite small, especially for the period before 1640, it would be beneficial to increase the number of sensationalist news pamphlets looked at when inventorying their textual labels, especially if one wants to identify any diachronic trends. This could be quite easily done by expanding the list of sensationalist topics to include murders, highway robberies, and other crimes as well as fires, earthquakes, floods, and other natural disasters. Alternatively, using the descriptive adjectives that often occurred with the textual labels (such as *most true* or *wonderful*) as title keywords would further increase the range of texts.

The second claim that this study tested is whether the design of the title pages of sensationalist news pamphlets shifted in the 1640s from title pages emphasizing the overall visual look of the page to those highlighting textual labels only (Suhr 2011: 125–142). This claim was also corroborated by the current study: it can indeed be said that the 1640s were a watershed in title page design. This shift suggests that a critical mass of readers and texts had been reached, such that readers were familiar enough with the generic conventions of sensationalist news pamphlets to recognize them on the basis of their textual labels rather than on the basis of the visual image of the title page. An important contributing factor to this was increased literacy, but also an increase in reading habits and a wider availability of texts, both of which may be attributed to the explosion of print and the desire for news in the turbulent 1640s. The intriguing rise in highlighting topical labels, which would be worth investigating further, may also be tied in with the dramatic increase in the availability of reading material, not only of sensationalist news pamphlets but also of periodical publications in the late 17th century.

6.   Conclusion

Historical linguists especially in the field of pragmatics have in recent years begun to take into account the many ways in which visual features of a text work together with its linguistic features, for example in structuring text and guiding readers' interpretations and comprehension. In my earlier work (2011: 142), I problematize the assumption that the presence of a textual label in the titles of texts aimed at semi-literate audiences would be automatically recognized by contemporary readers as such and suggest that it is only when textual labels are highlighted typographically that they can be considered reliable indicators of the recognition of textual labels as genre or text-type descriptors. This study supports this view by showing that typographical highlighting of textual labels was a general development in sensationalist news pamphlets from the 1640s onwards. Furthermore, it shows that certain textual labels were favored over others and that clustering them together with topical words was used as a strategy for specifying the rather general textual labels.

It would be interesting to find out when typographical highlighting of textual labels began to occur systematically in more literate genres. Bach's study (1997) indicates that visual highlighting of textual labels was considered important in polemical pamphlets of the 1640s and 1650s, but what of earlier polemical pamphlets? If such highlighting is absent in polemical pamphlets before 1640, it is possible that the 1640s watershed in title page design observed in sensationalist news pamphlets merely reflects a larger contemporary shift rather than developments in the functionality of textual labels for semi-literate audiences. On the other hand, earlier textual highlighting in more literate and established genres – polemical pamphlets but also other genres – could show that the conventionalization of textual labels took place at different paces for different audiences. A pilot study of plague texts in the Stuart era (Ratia 2013) points to this hypothesis, but much more work needs to be done with textual labels of historical genres.

References

Bach, Ulrich. 1997. *Englische Flugtexte im 17. Jahrhundert: Historisch-pragmatische Untersuchungen zur frühen Massenkommunikation* (Anglistische Forschungen 254). Heidelberg: Universitätsverlag C. Winter.

Bös, Birte. 2015. Conceptualisations, sources and agents of news: Key terms as signposts of changing journalistic practices. In Birte Bös & Lucia Kornexl (eds.), *Changing genre conventions in historical English news discourse*, 23–51. Amsterdam: John Benjamins. doi:10.1075/ahs.5.02bos

Brayman Hackell, Heidi. 2005. *Reading material in early modern England: Print, gender and literacy*. Cambridge: Cambridge University Press.

Brownlees, Nicholas. 2015. "We have in some former bookes told you": The significance of meta-text in 17th-century English news. In Birte Bös & Lucia Kornexl (eds.), *Changing genre conventions in historical English news discourse*, 23–51. Amsterdam: John Benjamins. doi:10.1075/ahs.5.01bro

Carroll, Ruth, Matti Peikola, Hanna Salmi, Mari-Liisa Varila, Matti Peikola & Risto Hiltunen. 2013. Pragmatics on the page: Visual text in late medieval English books. *European Journal of English Studies* 17(1). 54–71. doi:10.1080/13825577.2013.755006

Claridge, Claudia. 2000. Pamphlets and early newspapers: Political interaction in news reporting. In Friedrich Ungerer (ed.), *English media texts past and present: Language and textual structure* (Pragmatics and Beyond, n.s., 80), 25–43. Amsterdam: John Benjamins. doi:10.1075/pbns.80.04cla

Claridge, Claudia. 2003. "Life is ruled and governed by opinion": The Lampeter corpus of Early Modern English tracts manual of information. https://www.tu-chemnitz.de/phil/english/sections/linguist/real/independent/lampeter/manual/pages/contents.html (26 February, 2015.)

Clarke, Sandra. 1983. *The Elizabethan pamphleteers: Popular moralistic pamphlets 1580–1640*. London: Athlone Press.

Davies, Marie-Hélène. 1986. *Reflections of Renaissance England: Life, thought and religion mirrored in illustrated pamphlets 1535–1640* (Princeton Theological Monograph Series 1). Allison Park, PA: Pickwick Publications.

Fox, Adam. 2000. *Oral and literate culture in England 1500–1700* (Oxford Studies in Social History). Oxford: Clarendon Press.

Lampeter corpus of Early Modern English tracts. 1999. Compiled by Joseph Schmied, Claudia Claridge & Rainer Siemund. In Knut Hofland, Anne Lindebjerg & Jørn Thunestvedt (eds.), *ICAME collection of English language corpora*, 2nd edn., CD-ROM. Bergen: The HIT Centre, University of Bergen.

Luborsky, Ruth Samson. 1987. Connections and disconnections between images and texts: The case of secular Tudor book illustrations. *Word & Image* 3(1). 74–85. doi:10.1080/02666286.1987.10435367

Meurman-Solin, Anneli & Jukka Tyrkkö (eds.) 2013. *Principles and practices for the editing and annotation of diachronic data* (Studies in Variation, Contacts and Change 14). Helsinki: VARIENG. http://www.helsinki.fi/varieng/series/volumes/14/ (9 October, 2017.)

Moxon, Joseph. 1683. *Mechanick exercises: Or, the doctrine of handy-works. Applied to the compositors trade*, vol. 2. London: Printed for Joseph Moxon.

Ratia, Maura. 2013. Investigating genre through title-pages: Plague treatises of the Stuart period in focus. In Anneli Meurman-Solin & Jukka Tyrkkö (eds.), *Principles and practices for the editing and annotation of diachronic data* (Studies in Variation, Contacts and Change 14). Helsinki: VARIENG. http://www.helsinki.fi/varieng/series/volumes/14/ratia/ (9 October, 2017.)

Raymond, Joad. 1996. *The invention of the newspaper: English newsbooks 1641–1649*. Oxford: Clarendon Press.

Raymond, Joad. 2003. *Pamphlets and pamphleteering in early modern Britain* (Cambridge Studies in Early Modern British History). Cambridge: Cambridge University Press.

Salmi, Hanna. 2013. Visual presentation of English body and soul debates. *Studia Neophilologica* 86(suppl. 1). 144–156. doi:10.1080/00393274.2013.834106

Shevlin, Eleanor F. 1999. "To reconcile book and title, and make 'em kin to one another": The evolution of the title's contractual functions. *Book History* 2(1). 42–77. doi:10.1353/bh.1999.0011

Suhr, Carla. 2011. *Publishing for the masses: Early modern English witchcraft pamphlets* (Mémoires de la Société Néophilologique de Helsinki 83). Helsinki: Société Néophilologique.

Thomas, Keith. 1986. The meaning of literacy in early modern England. In Gerd Baumann (ed.), *The written word: Literacy in transition. Wolfson College Lectures*, 97–131. Oxford: Clarendon Press.

Tyrkkö, Jukka, Ville Marttila & Carla Suhr. 2013. The Culpeper Project: Digital editing of title-pages. In Anneli Meurman-Solin & Jukka Tyrkkö (eds.), *Principles and practices for the editing and annotation of diachronic data* (Studies in Variation, Contacts and Change 14). Helsinki: VARIENG. http://www.helsinki.fi/varieng/journal/volumes/14/tyrkko_marttila_suhr/ (9 October, 2017.)

Van Dijk, Teun A. 1980. *Macrostructure: An interdisciplinary study of global structures in discourse, interaction, and cognition*. Hillsdale, NJ: Erlbaum.

Watt, Tessa. 1991. *Cheap print and popular piety, 1550–1640* (Cambridge Studies in Early Modern British History). Cambridge: Cambridge University Press.

"… all spirits, and are melted into air, into thin air"

Metaphorical connections in the history of English

Marc Alexander and Christian Kay

University of Glasgow

This article uses data from *Mapping metaphor with the Historical thesaurus* to describe conceptual connections between the supernatural and other areas of the lexicon across the history of English. By discussing the cases of angels and evil in some detail, we also argue that the data and worldview presented here about the spirit world challenge the conventional concrete/abstract division found in metaphor theory, and propose an alternative.

Keywords: metaphor, English, supernatural, Mapping metaphor, Historical thesaurus, angels, evil

0. Christian Kay, 1940–2016

Christian J. Kay, Emerita Professor of English Language at the University of Glasgow, died peacefully on 4 June 2016, aged 76. Over the course of her long career she was an exceptional and generous scholar and a good and close friend to colleagues worldwide in English historical linguistics. Her lifetime of work on the *Historical thesaurus of English* began in 1969; she saw its first edition (titled by the publisher as the *Historical thesaurus of the Oxford English dictionary*) published 40 years later to popular and scholarly acclaim. Despite retiring in 2005, Christian remained highly active as a researcher, including acting as co-investigator on major externally funded grants until 2015. This included work on a number of *Historical thesaurus* daughter projects, including the *Mapping metaphor* project we use in this article. She will be remembered by her friends and colleagues as a warm, incisive, and selfless academic and a researcher of distinction. Both authors presented this research jointly at the 2014 ICEHL conference in Leuven, and both approved the pre-final version of this article before Kay passed away. Any errors or infelicities introduced in the final review or proofreading phases are Alexander's alone.

1.   Mapping metaphorical connections

The process of metaphor is by now firmly established as a key lexical factor in linguistic change. If, therefore, somewhere between 8% and 18% of English discourse is metaphorical (as argued by Steen et al. 2010: 765), then the means by which lexical items extend themselves metaphorically and by which these new meanings become semanticized are particularly rich for historical linguistics (well summarized in Brinton & Traugott 2005). However, as data-driven linguists we are crippled by the lack of detailed, systematic information about how and where this process has occurred in the history of English. With the exception of detailed subject-specific surveys such as Allan (2009), the process of metaphor in the history of English has been investigated through the collection of what examples were available to a linguist. While understandable and altogether valid, this lack of systematicity is problematic, primarily because the key insight of the work done on corpus historical linguistics is that linguistic processes often look very different when we aim to account for large bodies of data rather than isolated examples; as John Sinclair famously said, "Language looks rather different when you look at a lot of it at once" (1991: 100).

We use data from the *Mapping metaphor with the Historical thesaurus* project (*MM*), which aims to look at all of the history of metaphor in English at once.[1] In particular, we demonstrate the use of both *Mapping metaphor* and the *Historical thesaurus of English* (*HT*) to challenge aspects of metaphor theory by looking in more detail at *MM* category THE SUPERNATURAL (1Q); this also allows us to raise some issues especially relevant to historical linguistics. In more detail, we show that the domain of the supernatural – and especially supernatural places and beings, such as *heaven* and *demons* – poses fundamental questions for metaphor study. As supernatural concepts are variously conceived as being literal and concrete or fanciful and abstract (or yet more complex combinations thereof), are deeply held, and change wildly across time, place, and culture, we argue they challenge fundamentally any simple view of metaphor as linking the simply concrete with the simply abstract.

The article is structured as follows. We introduce our methodology and *MM*'s process of analysis in the next section, then overview the supernatural data under analysis. We then give two case studies, turning first to *angels* and *heaven*, and then

---

to *elves, imps*, and *devils*, and in both we aim to show how the detailed data-driven analysis which the *MM* project uses gives rise to complex questions as to the nature of metaphor. We conclude by rethinking the nature of 'concreteness' in metaphor.

2.   Examining metaphor through lexical overlap

To examine the history of metaphor, we divided the 225,131 semantic categories of the *HT*[2] into 415 broader clusters, representing higher-order concepts such as 1G01 FOOD AND EATING, or 3D02 RULE AND GOVERNMENT. The codes for these clusters – which *MM* treats as categories – reflect the *HT* by being ordered into the physical world (beginning with 1), the mental world (beginning with 2), and the social world (beginning with 3), with the following letters and numbers reflecting their order in the *HT* dataset. In order to then see the metaphorical links between these categories, the *MM* project used 'lexical overlap', the phenomenon whereby metaphorically semanticized lexemes in one area of the vocabulary of a language carry with them the lexical form of their origin. For example, the lexemes *overshine*, *dazzle*, and *outblaze*, used to indicate success in an endeavor, are also found in the vocabulary for light; the overlap of lexemes between these two concepts therefore indicates the presence of a potential metaphorical extension between them.

The insight that lexical overlap indicates metaphoricity is not new (see, for example, Smith 1996), but the *MM* project was the first to *systematically* analyze overlap across 1,300 years of the history of English and to categorize the metaphors thus revealed using cross-coder verification techniques. For each *MM* category – for example, 1J25 LIGHT – the complete set of *HT* lexemes inside the category was searched against the remainder of the *HT* (that is, all of the *HT* with the exception of the 1J25 category). This operation returned huge numbers of overlapping lexemes, and through detailed manual analysis – by an initial coder, by a second coder checking this data, and then by a third coder reconciling multiple directions of the data; for example, the metaphor connecting 1J25 LIGHT to 2C01 GOOD was coded once by a coder examining overlap from 1J25 and again separately by one considering overlap starting from 2C01. Much overlap was not metaphorically motivated but was instead the result of homonymy, meaning that this detailed coding stage was essential to the quality of the resulting *MM* data. The result of this work, which involved manually analyzing over four million overlap lexemes, is a database of approximately 16,000 metaphorical connections, with the first cited date of that

---

2.   This article uses version 4.2 of the *HT*, first published online in 2015. See http://www.glasgow.ac.uk/thesaurus/ for more detail.

metaphor alongside sample lexemes and an indication of metaphor directionality, all freely available for use and analysis at www.glasgow.ac.uk/metaphor. The material is presented in a variety of formats, including visualizations in 'metaphor maps' showing the range of overlaps.

3. 1Q The supernatural

In order to demonstrate the use of both *MM* and *HT* results and to show some the issues it raises for historical linguistics, let us now examine in more detail the *MM* category 1Q THE SUPERNATURAL.[3] The category occurs at the end of the first of the three primary divisions of *HT*, 01 THE EXTERNAL WORLD; organized religion, on the other hand, occurs in the third primary division, 03 SOCIETY, as does 3H FAITH. This division reflects an early decision to separate the two on the grounds that the first belongs to the physical world in which we live, the second to the constructs of the social world.[4] In some ways, however, concepts associated with the supernatural and organized religion fulfill similar functions in the thinking of human beings in that both attempt to explain the apparently inexplicable by reference to forces outside the normal range of human perception and experience.

Category 1Q contains general words for supernatural beings, events, practices, conditions and so forth, and sub-sections for the more specific concepts of DEITY, ANGEL, DEVIL, HEAVEN, and HELL. There are 17,249 lexical items (words and phrases) identified by the computer program as occurring in 1Q and at least one other *MM* category; of these, 6,077 are unique word forms. Since the data is derived mainly from the *Oxford English dictionary* (*OED*), it includes Old English (OE) words which continued to be recorded after about 1150 but not those which had disappeared by that date. The complete recorded OE vocabulary is presented in a separate part of the *MM* project.

Not all of these overlaps are of interest, since they include homonyms and polysemies which have been picked up because of their forms rather than any semantic connection. Thus a link has been recorded between 1Q03 ANGEL and 3L02 MONEY on the basis of the occurrence of *angel*, *angelet*, and *seraph* in both. *Angel* and *angelet* are English coins (with a device of the archangel Michael but no metaphorical implications), while *seraph* is a homonym denoting a Turkish coin. All tokens have

---

3. For those using the print version of *HT* (version 1.0; Kay et al. 2009), 01.07 THE SUPERNATURAL contains the material discussed here. In the revised version 4.2 data, it occurs in 01.17 THE SUPERNATURAL. For the purposes of *MM*, the categories occurring there have been aggregated into a single category 1Q. Likewise, 03.01 FAITH is 3H FAITH in *MM*.

4. On the principles behind the original classification, see Kay (2010).

to be carefully considered, however, especially by the historical linguist, since they may include words or meanings which are now obsolete, either literally or metaphorically or both, but have had metaphorical significance in the past. The word *maumet* or *mammet*, for example, is recorded in English from c1225 until 1647 in the meaning "A false god; an image of a false god, an idol" (*OED mammet* noun 1a) and from there develops the metaphorical meaning "A person who is the tool or puppet of another; a man of straw", current from c1390 to 1593 (*OED* noun 2). Such a link might be of interest to cultural historians as well as to historical linguists.

Table 1. Links between 1Q and other *MM* categories

1Q THE SUPERNATURAL: overlapping categories
Total: 1471 links
Coded as metaphor: 331 (23% of total overlap)
Strong metaphor: 100 (30% of metaphorical categories; 7% of total)
Weak metaphor: 231 (70% of metaphorical categories; 16% of total)

The primary focus of the *MM* project, however, is on the strong conceptual links evidenced between those concepts represented by the *MM* categories. Examining those lexical items which realize these categories enables us to identify such strong conceptual connections; Table 1 shows that 1Q has some degree of overlap with 1471 categories elsewhere in *HT*. Some clues as to the strength of these links are then available statistically – for example, there are high proportions of shared lexemes between 1Q categories and those of 1I12 SIGHT, 1K01 EXISTENCE, and 2A08 IMAGINATION, revolving around the concept of ghostly presences. This quantification is the point at which human coders take over and analyze the lexical items in each category in order to decide whether a link is metaphorical or not. In the case of 1Q, 331 categories were coded as metaphorical. Of these, 231 categories were coded as weak metaphors, ones where a metaphorical link is present but is not strongly lexicalized and therefore not suggestive of conceptual metaphor. An example comes from 3L02 MONEY where *angel* appears in the original *HT* in a sub-category CAPITALIST/FINANCIER with reference, presumably, to both the power and the benevolence of angels, alongside *financial wizard*, again with supernatural links through *wizard* and *venture capitalist*, unconnected to this metaphor. The remaining 100 links were classified as strong metaphors, on the grounds that the categories in question contained items which were sufficiently numerous and semantically coherent to suggest the presence of conceptual metaphors that reveal systematic ways of thinking about and understanding the world. In other words, for at least some period in the history of English, these categories contain examples of the 'metaphors we live by' referred to in the title of Lakoff & Johnson's seminal book (2003 [1980]).

4. Angels real and imagined

Metaphor depends on the perception of some kind of similarity between source and target domains. A first scan of metaphorical links from 1Q to other *MM* categories reveals two tendencies in the perceived similarities: (i) the links tend to polarize between negative and positive; and (ii) the negative metaphors far exceed the positive.

Salient features of items in the supernatural domain which are the source of positive metaphors in other domains include appearance, personality, and behavior. For example, in 1D05 BABY AND YOUNG PERSON, we find a child referred to as an *angelet* and a *cherub*, who can also be *cherubic* and *cherubimical*.[5] 2B12 PHYSICAL APPEARANCE/BEAUTY yields *angel*, *cherub*, *seraph*, *seraphic*, and *peri* as terms for attractive people. Other salient features of items in the source domain also contribute to some metaphorical meanings in 2B12, such as "small size" in the case of *fairy* and "magic" in *charming* and *enchanting*. 3F04 VIRTUE focuses on another aspect of a similar range of words, yielding *angel*, *cherub*, and *seraph* with reference to virtuous, and especially innocent, people. In 2D08 LOVE AND FRIENDSHIP, *cherub*, *cherubin*, and *seraph* are recorded as terms for loved ones.

At this point we should consider a question that inevitably arises when examining metaphor in this area of vocabulary: how can we find characteristics in human beings similar to those in entities which either do not exist or which exist in a different world from our own, and how can this situation be accommodated within a theory of metaphor? The usual pathway of metaphor is from concrete to abstract: we generally conceptualize abstractions in terms of familiar objects such as parts of the body, as in the *head of the family* or the *strong arm of the law*. There is, however, nothing concrete about angels or any of the other sources of metaphor identified in 1Q, and coding such metaphors raised problems for the *MM* team. An early focus of discussion was the concept of HEAVEN, which occurs from Old English onwards in the meanings of "The expanse in which the sun, moon, and stars are seen, (esp. in earlier use) regarded as having the appearance of a vast vault arched over the earth; the sky, the firmament" (*OED* 1a, *MM* 1A20) and "In the Christian tradition (and hence more widely): the abode of God and of the angels and persons who enjoy God's presence, traditionally regarded as being beyond the sky; the final abode of the redeemed after their life on earth; a state or condition of being or living with God after death; everlasting life. Opposed to *hell*" (*OED* 5a, *MM* 1Q05). The first of these senses is clearly literal, but whether the second is concrete or abstract, literal or metaphorical, is a matter of faith rather than linguistics. If concrete or literal, then *OED* 6a, *MM* 2D05 "A place likened or compared to

---

5. On metaphors for children, see Wild (2010, especially 297–298).

heaven; a place of supreme bliss" derives from it by a normal metaphorical process; otherwise we have an abstraction deriving from an abstraction.[6]

Kövecses acknowledges this problem when discussing RELIGION as a common target domain. Target domains, he writes, "are abstract, diffuse, and lack clear delineation; as a result, they 'cry out' for metaphorical conceptualization" (2010: 23). He goes on:

> God, similar to the concepts of society and nation, is conceptualized as a person: Father, Shepherd, King, and the like. It follows from the metaphor that believers are viewed as God's children, sheep, or subjects. Other aspects of religious experience involve the conceptualization of such notions as eternity, life after and before death, and so on, *which are necessarily metaphorical, since we have no experience of them*.
> (2010: 26; our emphasis)

Such an ontological view of metaphor is widely accepted but may be questioned in the context of a diachronic study. A basic tenet of cognitive semantics is a recognition of the importance of cultural context in elucidating meaning. As Taylor writes, "Word meanings are cognitive structures, embedded in patterns of knowledge and belief; the context against which meanings are characterized extends beyond the language system as such" (Taylor 2003: 87). In historical semantics, it seems reasonable to suggest that this context should include the knowledge and beliefs of earlier users of the language. The fact that HEAVEN as a concept developed when English speakers were largely Christian in outlook, and believed in the reality of heaven and hell, perhaps enables us to categorize *heaven* (5a) as a literal source of metaphor rather than a target; we are saying something like "x is like y *as y was understood in the culture of the time*" – otherwise, without this specification, those speakers who had no belief in a literal heaven in the sky and also did not acknowledge this historical dimension of metaphor would therefore have no connection at all between the sky and the literal location of a physical afterlife; with no ability to bring in a conceptual link between them, *OED* senses (1a) and (5a) of *heaven* would be mere homonyms, with no conceptual relationship.

In order to understand other exponents of the metaphor represented by (6a), such as *heaven of heavens*, *third heaven*, and *seventh heaven*, we also need some knowledge of early cosmography (Jewish and Islamic as well as Christian) where *heaven* refers to "each of the 'spheres' or spherical shells, lying above or outside each other, into which astronomers and cosmographers formerly divided the realms of space around the earth" (*OED* noun 4).

---

6.   There are, of course, numerous other senses and sub-senses of *heaven* in both *OED* and *HT*, but we lack the space to discuss them here.

Another interesting contribution to the medieval worldview is the idea of 'natural magic' or 'magic natural', defined by the *OED* as "magic involving the manipulation of supposed occult properties of the natural world (usually excluding the conjuration of personal spirits); this skill as a subject of study" (*magic* noun 1b), recorded from 1387 to 1605 and in historical use thereafter. The phrase also occurs in the *Middle English dictionary* (*MED*), defined as "The knowledge of hidden natural forces (e.g. magnetism, stellar influence), and the art of using these in calculating future events, curing disease, etc." (*magik(e)*, noun A). While natural magic was essentially a practical field of study, focused on areas such as treating disease and forecasting events, its importance in medieval thinking helps to explain the contemporary readiness to attribute inexplicable events to supernatural causes.

An extreme example is the behavior of the incubus, defined as "A feigned evil spirit or demon (originating in personified representations of the nightmare) supposed to descend upon persons in their sleep, and especially to seek carnal intercourse with women. In the Middle Ages, their existence was recognized by the ecclesiastical and civil law" (*OED* noun 1, first recorded c1275 – this recognition comes in part from the authorization of witch-hunt manuals; see Mackay 2009: 126). Although this spirit is described by *OED* as "feigned", the suggestion that the existence of these evil spirits was recognized in law suggests that this meaning should be taken as literal in the medieval context in the same way as we have suggested above for *heaven* (5a). Note that the word *nightmare* has a similar origin in "A female spirit or monster supposed to settle on and produce a feeling of suffocation in a sleeping person or animal" (*OED* 1, first recorded c1300).

Incubus has two metaphorical meanings in *OED*: "2. A feeling of oppression during sleep, as of some heavy weight on the chest and stomach; the nightmare", first recorded in 1561; and "3. A person or thing that weighs upon and oppresses like a nightmare", recorded from 1648. Thus the quasi-medical explanation of discomfort (and other things!) experienced during sleep later becomes a more general metaphor for anything experienced as oppressive.

Nightmare follows a somewhat similar path within *OED* sense 2, moving from:

a. "Originally (usu. with *the*): a feeling of suffocation or great distress experienced during sleep. Now usually: a bad dream producing these or similar sensations; an oppressive, frightening, or unpleasant dream", first recorded in 1562

to:

b. "*fig.* and in extended use. An oppressive fear; a frightening experience or thing; a source of fear or anxiety", first recorded in 1834

and:

c. "*colloq.* In weakened use: a person, thing, or situation that is very difficult or frustrating to deal with; an unpleasant or bad experience or prospect; a catalogue of disasters", first recorded in 1904.

The adjectival use, "Having the quality of a nightmare; extremely distressing, frightening, or oppressive; nightmarish. Later in weakened use: terrible, awful, fraught with difficulty", occurs somewhat earlier, from a1821. *Mare*, the etymon of *mare* in *nightmare*, occurs in many Germanic languages and is attested in English from Old English until the early 17th century in the meaning "A spirit believed to produce a feeling of suffocation in a sleeping person or animal; a feeling of suffocation experienced during sleep; an oppressive or terrifying dream" (*OED mare*[2], 1a). Thereafter, it is recorded only in dialect or as an abbreviation of *nightmare*; it does not seem to have developed a further metaphorical use. An interesting, if minor, example of folk etymology occurs in *night-horse*, recorded twice in the 20th century as a synonym of *nightmare*, presumably with reference to *OED mare*[1] "a horse". A similar connection possibly occurs in a rare example of *hag* meaning "nightmare" in 1696: "It is to prevent the Night-Mare (viz.) the Hag from riding their Horses" (*OED hag*[1], 1c).

For present-day speakers of English, *nightmare* in both its literal and metaphorical meanings is far more common than *incubus*, with *nightmare* occurring 14 times more frequently than *incubus* in the *British national corpus* (*BNC*) and 93 times more frequently in the *Corpus of contemporary American English* (*COCA*). However, the *OED* dates suggest that both words (and *mare*) were known in their first meanings to speakers of Middle English and that both literal and metaphorical meanings were known from Early Modern English. The overlap of dates also suggests that these earlier speakers could have had an understanding of the link between source and target concepts which is lost to many or most present-day speakers. Traditionally, metaphors where the source is no longer recognized are known as 'dead' metaphors, but discussion following Lakoff & Johnson (2003 [1980]) and further proposals in Lakoff (1987) have led to the use of the term 'historical metaphor' for cases where a conceptual metaphor survives even though its presence may not be apparent in all of its exponents. One of Lakoff's examples, discussed more fully in Allan (2013), is UNDERSTANDING IS HOLDING, represented in modern English by *grasp*, where the metaphorical link is obvious, and *comprehend*, where the link may be appreciated only by Latin scholars. Following this view, we can say that *incubus* and *nightmare* are linguistic exponents of a conceptual metaphor linking bad spirits and bad events, which, as we see in Section 5, persists throughout the history of English regardless of the status of individual members of the metaphor for any particular group of speakers.

5.   Roots of evil

Supernatural creatures of malign character are a fruitful source of metaphors for wicked people and evil circumstances. Strong metaphors based on this link occur in a wide range of *MM* categories including HARM (1O17), ADVERSITY (1O18), BAD BEHAVIOUR (1O22), DISORDER (1P15), DECEPTION (2A22), HEALTH (both physical and mental; 1C01–1C03), APPEARANCE (ugliness; 2B12), MORAL EVIL (3F05), and EMOTION (e.g. suffering, fear, anger, hatred; 2D01–2D17). An example is the behavior of the 'little people' – elves, fairies, imps, goblins, pixies, and so on. OE *ælf*, for example, is defined literally as "elf" in the *Dictionary of Old English A–G* (*DOE*) and then as "the supposed agency in illness or disease", followed by the phrase "*ylfa gescot* 'elf-shot', a medical condition of unspecified nature (apparently understood as a supernatural assault leading to illness or disease)".[7] *Elf-shot* is recorded in *OED* in 1681 and 1841, though given the time-gap, there may not be a direct link with the OE phrase. OE compounds include *ælf-adl* "elf-disease (of uncertain nature)", *ælf-cynn* "race of elves, referring to their supposed agency in bringing about some affliction; or perhaps referring to the affliction itself", *ælf-sogeþa* "disease thought to have been caused by supernatural agency, perhaps anaemia", and *ælf-siden* "influence or enchantment of elves, referring to an affliction of uncertain identity thought to be caused by supernatural agency and attended by fever" (all definitions from *DOE*). In later periods, elves could be either malevolent or benevolent in their dealings with humankind, and the first metaphorical meaning in *OED* is "A tricksy, mischievous, sometimes a spiteful and malicious creature" (*elf*[1], 2b, recorded a1556–1820). Fairies are likewise ambivalent in intention, being "a class of supernatural beings having human form, to whom are traditionally attributed magical powers and who are thought to interfere in human affairs (with either good or evil intent)" (*OED* 3a). In later use, the focus tends to be on the more attractive characteristics of fairies, such as their prettiness and small size, and a distinction is often made in traditional tales like *Sleeping Beauty* between *good fairies* and *wicked fairies*.

One of the most interesting words referring to small mischievous creatures is *imp*. Its somewhat eccentric semantic history can be summarized as follows, using the information in *OED*:

> 1a.   A young shoot of a plant or tree (OE–1672)
> 1b.   Applied to people (1377, 1596)
> 2a.   A shoot used in grafting, a graft (1377–1706)
> 2b.   Applied to people (1583–1615)

---

7.   Cf. German *Hexenschuss* "lumbago" and OE *hægtessan gesecot* (with thanks to Hans-Jürgen Diller).

3a. Scion (esp. of a noble house); offspring, child (usually male) (c1412–1611, then archaic) [e.g. 1548 *Hall's Vnion: Henry VIII* f. ccxlijv, That his sonne prince Edward, that goodly ympe, maie long reigne ouer you.]

3b. Used figuratively in the same way as *child* (c1380–1809) [e.g. *imp of fame*]

4. A 'child' of the devil or of hell

4a. With parentage expressed: Applied to wicked men, and to petty fiends or evil spirits (1526–1821) [e.g. *imp of Satan*]

4b. A little devil or demon, an evil spirit (1584–1882)

4c. Applied to people, often humorously (1633–1857)

5. A mischievous child (having a little of 'the devil' in him); a young urchin: often used playfully (1642–1859)

6. A young man, youth (1578–1648, then archaic).

The verb *imp*, also attested from Old English onwards, has a much more straightforward history, with a transferred sense of engrafting feathers in falconry and figurative uses based on the idea of implanting something.

Leaving aside senses 2 and 6, the mappings involved in the development of this set of meanings can be listed as follows:

Source: PLANTS	Target: PEOPLE/CREATURES	
1. Shoot/offspring of plant	3.	Offspring of humans
	4a.	Offspring of supernatural evil creatures
	4b.	Supernatural evil creatures
	4c.	Human evil creatures
	5.	Young human creatures

The category PLANTS is one of the common source domains for metaphors listed by Kövecses (2010: 19), and often occurs in the conceptual metaphor PEOPLE ARE PLANTS. The mapping from (1) to (3) above is a clear example of this metaphor, which can be formulated as OFFSPRING OF PEOPLE ARE OFFSPRING OF PLANTS. (4a) could perhaps be seen as an extension of this metaphor if we incorporate through blending a second metaphor SUPERNATURAL EVIL CREATURES ARE PEOPLE, which then takes over in (4b), largely losing the "offspring" element, which occurs in only one of the seven *OED* citations. A new metaphor with reversed directionality, EVIL HUMANS ARE SUPERNATURAL CREATURES, occurs in (4c) and in a weakened meaning in (5), where the notion of "child" reappears, though probably without any reference to the original metaphor.

The example of *imp* illustrates some of the problems involved in tracing the development of historical metaphors, where there may well not be a straightforward series of links in the diachronic profile of a polysemous word. In other cases, connections may be more obvious. *Devil*, for example, is also recorded from Old English. Its meaning progresses from theological senses to the more general literal

sense of a malevolent supernatural creature (*OED* 3) to wicked humans (*OED* 4a) or merely mischievous ones (4b), and is then, in the language of the first edition of the *OED*, "Applied in contempt or pity (chiefly with *poor*): A poor wretched fellow, one in a sorry plight, a luckless wight" (4c). By metonymy, it is also applied to the qualities of devils and people who resemble them (*OED* 6).

Metaphors attributing evil supernatural qualities to people often result in semantic bleaching, as is the case with both *imp* and *devil* and their derived forms, as well as many other words. To describe something as *diabolical* or *fiendish* nowadays implies fairly mild criticism, and *impish* can even be complimentary in collocations such as *impish laughter* or *smiles*. Such amelioration of meaning may result from a natural tendency to overstatement; in the case of adverbs, bleaching is part of a process of grammaticalization leading to the use of words originally referring to supernatural or human evil as intensifiers in expressions such as *fiendishly busy* (1879), *devilishly pretty* (1782), and *devilish glad* (1807). Similar bleaching processes occur in many adverbs which have lost semantic values and act as general intensifiers (*awesomely, dreadfully, frightfully*, and so on; see also Kay & Allan 2015: 82ff.). Alternatively, the weakening of meaning may indicate a loss of belief in such creatures in modern culture (see Wilson & Huff 2001; Baker 2008) and hence a loss of the need to take them seriously.

6.   Conclusion

Overall, the spirit world as represented in English metaphor is used to explain the evils and misfortunes of humankind rather more than its virtues and successes. There is also sometimes an attempt to evade responsibility for events by attributing misfortunes such as an incubus-induced pregnancy or unexplained illnesses to a supernatural agency. The data presented here indicates some ways in which the alleged influence of the supernatural world can be used as excuses or rationalizations of unfortunate or malign events in the physical world. In a broader way, however, the data also makes manifest the importance of not applying currently dominant worldviews to the history of English.

It is tempting to characterize the worldview of the past as one which contrasts concrete and abstract as we do, despite the fact that this contrast is one born primarily of the 20th century (see, amongst others, Goodman & Quine 1947). When confronted with the methodology of the *Mapping metaphor* project, this is particularly stark, given that the project focuses on rigorously accounting for substantial amounts of data and so building an account of metaphor in the history of English from the small to the large scale. Given the data above, we propose that in this case, rather than contrasting concrete and abstract or literal and figurative, it is more

useful to invoke *at the least* a tripartite worldview, which includes a literal world, a metaphorical world which is accepted as literal in some belief systems, and a wholly metaphorical world.

Such a view accords with that put forward by Kövecses (2006: 181–204). While maintaining his position that all language used to express abstract ideas must initially have been a metaphorical transfer from the physical, he allows that "A part of our conceptual system consists of abstract concepts that are metaphorically defined. The definition of abstract concepts by means of metaphor takes place automatically and unconsciously" (2006: 201). His examples of such metaphorically defined concepts include EMOTIONS ARE FORCES and ABSTRACT COMPLEX SYSTEMS ARE BUILDINGS. In such cases, the initial metaphorical mapping is so basic to our thinking that it can be taken for granted and treated as literal and so its entailments can generate further metaphors.

The recategorization of literal and metaphorical involved in this approach offers some support to the solutions proposed above for the problems involved in analyzing metaphors in the supernatural domain. However, it does not entirely go far enough; as we show above, the *MM* data indicates that any approach to metaphor which aims to make distinctions between clusters of domains (such as in the identification of a separate domain complex called the 'abstract') must take into account the tendency for language to form fuzzy sets. While it is possible in a generalization to refer to 'concrete' and 'abstract' as distinct sets, data-driven and context-aware analyses such as that undertaken here will at times need to form from the fuzzy and overlapping boundary a range of other intermediate categories. We tentatively propose four of these: firstly, 'culturally concrete' (described earlier in this article); secondly, 'idiolectally concrete', which covers ideas which are concrete for only some persons (such as someone who might believe in particular types of supernatural being within a wider cultural and linguistic environment which does not accept their existence); thirdly, instances of agreed social, political, scientific, moral, or ethical approximations or conventions (where persons conventionally act as if something is real, even though it is known not to be); and finally, various combinations or complexes of these categories. A new type of analysis is also needed of the occasions where metaphors and other conceptual connections historically transition from one state to another, representing a shift in worldview.

All of this is to say that, as ever, the need to account rigorously for complex real-world data leads to significant theoretical insights. In addition to the argument we present above, we would also strongly aver that the rich array of data now available from the *Mapping metaphor* project enables researchers to take such investigations yet further, and develop fresh understandings of the place of metaphor in our world.

References

Dictionaries and online resources

BNC = Davies, Mark. 2004–. *British national corpus (BYU–BNC).* Based on the British national corpus from Oxford University Press. http://corpus.byu.edu/bnc/

COCA = Davies, Mark. 2008–. *The Corpus of contemporary American English: 450 million words, 1990–present.* http://corpus.byu.edu/coca/

DOE = Cameron, Angus, Ashley Crandell Amos & Antonette diPaolo Healey (eds.). 2008. *Dictionary of Old English, A–G.* Toronto: Pontifical Institute of Mediaeval Studies. http://www.doe.utoronto.ca

HT = Kay, Christian, Jane Roberts, Michael Samuels, Irené Wotherspoon & Marc Alexander (eds.). 2015–. *Historical thesaurus of English, version 4.2* Glasgow: University of Glasgow. http://www.glasgow.ac.uk/thesaurus

MM = 2015–. *Metaphor map of English.* Glasgow: University of Glasgow. http://www.glasgow.ac.uk/metaphor

MED = Kurath, Hans & Sherman M. Kuhn (eds.). 1956–2001. *Middle English dictionary.* Ann Arbor: University of Michigan Press. http://quod.lib.umich.edu/m/med/

OED = Murray, Sir James A. H., Henry Bradley, Sir William A. Craigie & Charles T. Onions (eds.). 1884–1928. *Oxford English dictionary*; Supplement and bibliography 1933. Supplement, 1972–1986; Robert W. Burchfield (ed.). 2nd edn., 1989; John A. Simpson & Edmund S. C. Weiner (eds.). Additions series, 1993–1997; John A. Simpson, Edmund S.C. Weiner & Michael Proffitt (eds.). Oxford: Oxford University Press. 3rd edn. (in progress). OED online, March 2000–, John A. Simpson (ed.) (2000–2013), Michael Proffitt (2013–). www.oed.com

Books and articles

Allan, Kathryn. 2009. *Metaphor and metonymy: A diachronic approach.* (Publications of the Philological Society 42). Chichester: Wiley-Blackwell.

Allan, Kathryn. 2013. An inquest into metaphor death: Exploring the loss of literal senses of conceptual metaphors. *Cognitive Semiotics* 5(1–2). 291–311.

Baker, Joseph. 2008. Who believes in religious evil? An investigation of sociological patterns of belief in Satan, hell, and demons. *Review of Religious Research* 50(2). 206–220.

Brinton, Laurel J. & Elizabeth Closs Traugott. 2005. *Lexicalization and language change.* Cambridge: Cambridge University Press. doi:10.1017/CBO9780511615962

Goodman, Nelson & Willard V. O. Quine. 1947. Steps toward a constructive nominalism. *Journal of Symbolic Logic* 12. 105–22. doi:10.2307/2266485

Kay, Christian. 2010. Classification: Principles and practice. In Michael Adams (ed.), *Cunning passages, contrived corridors: Unexpected essays in the history of lexicography*, 255–270. Monza: Polimetrica.

Kay, Christian & Kathryn Allan. 2015. *English historical semantics.* Edinburgh: Edinburgh University Press.

Kay, Christian, Jane Roberts, Michael Samuels & Irené Wotherspoon (eds). 2009. *Historical thesaurus of the Oxford English Dictionary*, 2 vols. Oxford: Oxford University Press.

Kövecses, Zoltán. 2006. *Language, mind, and culture.* Oxford: Oxford University Press.

Kövecses, Zoltán. 2010. *A practical introduction to metaphor*, 2nd edn. New York: Oxford University Press.

Lakoff, George. 1987. The death of dead metaphor. *Metaphor and Symbolic Activity* 2(2). 143–147. doi:10.1207/s15327868ms0202_5

Lakoff, George & Mark Johnson. 2003 [1980]. *Metaphors we live by*, 2nd edn. Chicago: The University of Chicago Press. doi:10.7208/chicago/9780226470993.001.0001

Mackay, Christopher S. 2009. *The hammer of witches: A complete translation of the Malleus Maleficarum*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511626746

Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Smith, Jeremy J. 1996. *An historical study of English: Function, form, and change*. London: Routledge.

Steen, Gerard J., Aletta G. Dorst, J. Berenike Herrmann, Anna A. Kaal & Tina Krennmayr. 2010. Metaphor in usage. *Cognitive Linguistics* 21(4). 765–796. doi:10.1515/cogl.2010.024

Taylor, John. 2003. *Linguistic categorization*, 3rd edn. Oxford: Oxford University Press.

Wild, Kate. 2010. Angelets, trudgeons and bratlings: The lexicalization of childhood in the *Historical thesaurus of the Oxford English dictionary*. In Michael Adams (ed.), *Cunning passages, contrived corridors: Unexpected essays in the history of lexicography*, 289–308. Monza: Polimetrica.

Wilson, Keith M. & Jennifer L. Huff. 2001. Scaling Satan. *The Journal of Psychology* 135(3). 292–300. doi:10.1080/00223980109603699

Historical layers in text and genre

Conservatism and innovation in Anglo-Saxon scribal practice

Christine Wallis

University of Sheffield

The text of the Old English *Bede* found in Cambridge, Corpus Christi College, MS. 41 (B) is remarkable for its extensively updated language, when compared with other, earlier *Bede* manuscripts. This paper compares B with other manuscripts of the *Bede* to uncover some of the scribal decisions which shape the surviving text. B's text is subject to many alterations, indicating a translator scribe who frequently updated and altered the language of his exemplar (i.e. the manuscript from which he copied to produce the present text). However, the presence of a number of non-sensical readings points to a scribe who sometimes struggled to make sense of the text in front of him and whose abilities did not extend far enough to create a good reading in the face of these difficulties. These scribal decisions allow us to identify factors which influenced the shape of B's text, such as the interplay between B's now-lost exemplar and its scribes' working methods. Careful analysis of some of B's linguistic features enables us to draw conclusions about the age and status of its exemplar and to recover some part of a lost *Bede* manuscript.

Keywords: Old English, late West-Saxon, Mercian, philology, manuscripts, scribal practice, prose translations, Bede

1. Introduction

Of the four main surviving manuscripts of the Old English *Bede*, one in particular, Cambridge, Corpus Christi College, MS. 41 (B), has attracted attention for the way its scribes have updated the language of the text, turning the *Bede*'s original Mercian dialect into the late West-Saxon more familiar to its 11th-century scribes.[1] The

---

1. The other surviving witnesses of the Old English *Bede* are the following: Oxford, Bodleian Library, MS. Tanner 10 (T), s. x[1]; Oxford, Corpus Christi College, MS. 279B (O), s. xi[in]; Cambridge, University Library, MS. Kk 3.18 (Ca), s. xi[2]. In addition to these four complete (or near-complete) manuscripts, a badly fire-damaged copy exists in London, British Library, MS. Cotton Otho

manuscript dates from the first half of the 11th century, according to Ker (1957: 43), and was completed by two scribes (here referred to as B1 and B2). The first scribe (B1) wrote from the beginning of the manuscript as far as p. 190 (towards the end of Book 3 of the *Bede*), while the second (B2) was responsible for the text from p. 190 to the end. A third scribe copied a wealth of material into B's margins, but as the hand is not contemporary with the main text, and none of these marginal texts are part of the Old English *Bede*, this paper considers only the performances of the first two scribes.

While the manuscript gives us some clear evidence of the kinds of changes late-Anglo-Saxon scribes could (and did) make to the texts they copied, it is difficult to contextualize these changes in the case of B. We know that the manuscript was in the possession of Exeter cathedral by the second half of the 11th century, as a note recording its donation by Exeter's archbishop Leofric survives on p. 488. However, studies of other manuscripts donated by Leofric during his episcopacy (1046–1072) show that, while some books were produced in Exeter itself, others were acquired from elsewhere to furnish the cathedral (Treharne 2003, 2007). As B is not written in a hand that has been identified with Exeter, and as it predates Leofric's episcopacy, it seems likely that it is a manuscript which was produced in another center, before being brought to Exeter. Budny (1997: 507) suggests a minor center in the south of England as a probable place of origin, while Grant (1989: 8) tentatively posits a Winchester connection, based on a comparison of B's main and marginal texts with those found in other manuscripts. On the other hand, Stokes (2014: 142) notes that palaeographically B shares similarities with manuscripts for which he suggests an origin at Crediton (Oxford, Bodleian Library, Eng. hist. a.2; London, British Library, Stowe Charter 34). Although firm evidence for such an attribution is lacking, he states that "production at a house like Crediton seems entirely plausible."

B was rather neglected by early commentators. When preparing his edition of the *Bede*, Miller was disappointed with B's text because it deviated so strongly from the Mercian character of the earlier manuscripts. He abandoned his plan to base an edition on the manuscript because "the scribe or editor of B's text has dealt very freely with his author, changing forms and words and recasting sentences" (1890: v–vi), and instead he worked from T, whose more Mercian character he felt to be closer to the language of the *Bede*'s archetype. B's text was edited (along with that of O) by Schipper (1899), but in spite of this, comparatively little work focused on B itself until the late 20th century. More recent work by Grant (1989) and Rowley (2004, 2011) has done much to rehabilitate B's reputation as a manuscript worthy

---

B.xi (C), s. x^{med}, from which a transcript was made by Laurence Nowell in the 16th century (CN). Finally, London, British Library, MS. Cotton Domitian ix, f.11 (Z), s. x^{in} contains three short excerpts from the Bede. All dates are those given by Ker (1957).

of study. While Campbell (1951: 351) had noted that B was "the most radical and violent of all the manuscripts in its changes of the text," Grant (1989) focuses on the differences between B's West-Saxon language and the Mercian of earlier manuscripts. Rowley points out that it is this dialect shift itself which has prompted scholars to view B as "a mediocre text in the wrong language" (2004: 11), when compared with earlier texts containing more Mercian features, such as T.

This article consists of three main sections. After outlining the methodology used to uncover the scribal layers within the text, I show how this reveals some examples of the exemplar forms retained by B1. Following this, B1's behavior as a textual emender is examined, with a focus on the methods he used to overcome the textual problems he encountered. Despite his effective shaping of B's text in many ways, there are some notable instances where B1 fails to control the *Bede*'s text, resulting in emendations which do not appear to make sense. I discuss three groups of emendations. First are those which can be described as successful in that they produce text which, although altered from that of the exemplar, provides the reader with a credible reading and one in which a textual alteration would not be noticed without comparison with other manuscripts. The second group consists of emendations in which word division and spacing appear to play a role, while the third group includes those where the new reading seems to be dependent on the shape and sequence of the exemplar's graphs. This study is based on a close examination of the text of Book 3 of the *Bede*, as it appears in manuscripts T, O, Ca, and B. The majority of the selected text was copied by B1, with a smaller section by B2.[2]

2. Methodology

When a scribe copies a manuscript, there are three courses of action available to him. Firstly, he may copy the text exactly as it appears in his exemplar and produce, through his literatim copying, a manuscript that is identical to its parent. Alternatively, the scribe can 'translate' the exemplar text as he writes and update the original language into dialectal, grammatical, or lexical forms that accord with those of his own training or preferences. As a third option, the scribe may combine the first two approaches, updating some features, while preserving others (Laing 2004). These preserved, relict features will then show through the later layer of text,

---

2. Book 3 of B contains just under 18,000 graphic units. Of these, about 2,000 are by B2. All quotations from the *Bede* are presented as they appear in the relevant manuscript. References are to the page/folio of the manuscript, with additional references by page and line number to Miller's (1890) edition.

giving us a window onto the textual features of its exemplar.[3] Therefore we may find ourselves confronted with a surface text written at one point in time, in one geographical location, but in this text linguistic information may be relayed which is a relict of other times and/or places (Benskin & Laing 1981: 58–59). If later scribes copy the text into a different dialect, then one way in which these relict forms may present themselves is as linguistic forms which are inconsistent with the dialect of the latest scribe. In the case of the *Bede's* later manuscripts, these relict forms may surface as occasional Mercian features in a predominantly late West-Saxon text.

There are other ways of making distinctions between linguistic features in the *Bede*, besides discerning dialect differences between Mercian and late West-Saxon. For example, we can examine 'older' and 'more modern' features, and while diachronic change certainly accounts for some of the visible changes, other differences are more nuanced and may be due to the preferences or training of the latest scribe. It is perhaps more nuanced to think of several continua along which various features can move. While a chronological continuum could be established, which has the benefit of being applicable to texts other than the *Bede*, we could also use a continuum which is specific to the texts of the *Bede*, as in Figure 1:

Conservative ⟶ Innovative

Earlier OE features   Later OE features
Mercian   West-Saxon
Exemplar forms   Scribal innovation

Figure 1. The 'conservative–innovative continuum' for the *Bede*

We are able to propose this continuum because the four manuscripts examined in this study can be independently dated through palaeographic means,[4] and from this it is evident that certain Mercian features appear more frequently in the older texts. While also being aware that no manuscript is a direct copy of any other in this group, a chronological trend is discernible whereby earlier texts are more likely to contain more Mercian features and later texts are more likely to contain more West-Saxon features.

The *Bede*-specific continuum has been labeled 'conservative-innovative', where features are 'conservative' *in the context of the Bede* if they are (i) features found

---

3. See also the contribution by Minkova in this volume, where the impact of scribal training and spelling norms is discussed in relation to coin evidence.

4. See Stokes (2014: 68): "texts and even linguistic forms can be copied from one manuscript to another, making attribution difficult; but script is much harder to imitate and scribes generally had little reason for doing so."

more commonly in the earlier texts, (ii) features associated with a Mercian dialect, and (iii) features which (through a combination of (i) and (ii)) are attributable to a manuscript's exemplar. Having established a number of features which are either 'old' or 'Mercian', which are more commonly found in the earlier *Bede* manuscripts and which are found to a far lesser extent in the later manuscripts, we can use this group as a kind of measure against which other 'conservative' features can be compared. Innovative features, on the other hand, are those which appear to have been introduced into the *Bede* text by a scribe as a result of his own preferred forms, rather than being a relict of the exemplar.

Questions of the transmission of the Latin and Old English versions of the *Bede* have formed the basis of two recent studies. Lapidge (2008) has shown through a detailed comparison of the surviving Latin and Old English texts that the manuscript from which the Old English translation was made was a more faithful copy of Bede's Latin text than any of the surviving Latin witnesses. His study is particularly useful for us as he demonstrates "how an indirect witness can illuminate a transmissional history" (2008: 245), even when the intervening manuscripts have been lost. Waite (2014) explores the archetype of the Old English *Bede* by comparing the dialect vocabulary of different surviving manuscripts. He advocates the study of the translator's style, in particular the way in which certain Latin terms were translated into Old English. According to Waite, understanding the translator's systematic and precise method allows us to detect the places where lexical substitution is most likely to have been undertaken in subsequent witnesses. Both studies demonstrate fruitful lines of enquiry where exemplars are absent; however, they focus on what later transmission can reveal about lost archetypes and exemplars. The focus of this article, in contrast, is on what such a study can tell us about the motivations, methods, and practices of the scribes who produced our surviving witnesses.

There are several kinds of innovative features introduced by B's scribes. For example, in contrast with the earliest *Bede* manuscripts, B's text tends to contain spellings reflecting breaking before <l>+C rather than retraction (e.g. *eall* rather than *all*), <a> before nasals (e.g. *fram* rather than *from*), and <ode> spellings in weak class 2 preterite verbs (*leornode, willnode*, rather than *leornade, willnade*). Of all the *Bede* manuscripts, B is the most consistently West-Saxonized (Wallis 2013: 96–111). Therefore, B1's overall behavior can be categorized as that of a translator scribe. However, he retains a number of small but intriguing exceptions – relict forms – in his copy. That these textual oddities are relicts from the exemplar is highly likely, because the same kinds of features occur in the other *Bede* manuscripts, occasionally in the same place in the text. In addition to relict features, some of B1's innovative forms appear to be attempts to mend or improve on the exemplar text. This has some important implications for what we can deduce about the age and status of B's exemplar. In the next section we consider the relict forms in B.

3.   Relict forms

Several types of relict form appear throughout Book 3 of B. Relict forms are important evidence for the behavior of the scribes, as they provide a window onto the state of the exemplar. In B's case, comparison with the other surviving *Bede* manuscripts allows us to map a variety of relict forms present in the exemplar; moreover, it suggests that this exemplar lay at the conservative end of the Old English *Bede* continuum. B's relicts include palaeographic features, such as *f*-shaped <y>, as well as morphological and orthographical ones, such as denasalization and double vowels. What is important is that all these features occur in other *Bede* manuscripts, and especially in the earliest ones, and so they can be placed at the conservative end of the *Bede*-specific continuum outlined above.

3.1   *f*-shaped <y>

B1 has a tendency to write an *f*-shaped <y>, which according to Stokes was "abandoned quite early in the eleventh century" (2014: 190). Ker believed that in the case of B, this uncommon form of <y> was probably written in imitation of the exemplar: "this last form [*f*-shaped <y>], still common in s. x, is rare later and was sometimes used, no doubt, only because the scribe found it in his exemplar" (1957: xxxi). In contrast, B2 never writes *f*-shaped <y>, as a comparison of Figures 2 and 3 shows:[5]

Figure 2. B1's *f*-shaped <y> (*þy*, *bysceop*, and *bebyrged*), and undotted <y> (*swylce*) (p. 190)

Figure 3. dotted <y> in B2's hand (*ylce*, *swylce*, *gehwyllcum*, *cwylmde*) (p. 190)

Figures 2 and 3 also show that neither B1 nor B2 are consistent in the type of <y> they use; B2 frequently uses a dotted <y>, as well as an undotted variant which is shared by B1 in his section of the manuscript. These two illustrations both come from the same manuscript page, where B2 takes over from B1's stint, suggesting that the variation in <y> is indeed due to the preferences of each particular scribe rather than to a pattern in the underlying exemplar.

As a proposed relict form, *f*-shaped <y> is unusual in being a palaeographic feature. However, it also occurs occasionally in O, when written by its second scribe. Another palaeographic relict occurs, also in O, where the third scribe uses half-uncial <r> in his short stint (Wallis 2013: 71). As Ker and Stokes identify *f*-shaped <y> as an archaic form by the time B was written, its presence suggests an exemplar that was not new at the time B1 made his copy. If my reasoning is correct, then this older manuscript may have also contained several other conservative *Bede* features in its text, as the following sections show.

3.2   Denasalization

In at least three places in Book 3, B transmits a reading with denasalization. Denasalization is a feature which in the *Bede* particularly affects plural subjunctive verbs, whereby the verb loses its final <n>, resulting in readings such as *wolde* for *wolden*. Campbell notes the presence of denasalization in Northumbrian and Mercian texts (1959: Section 472), and it is found to varying degrees across all *Bede* manuscripts. One instance where denasalization is found in all four main manuscripts is given in (1):

(1)   *þa bæd he osweo þone cyning þ he him sume lareowas sealde. þa þe his ðeode*
      to cristes geleafan **gecyrde** 7 *mid þam halwendan wylle fullwihtes baðe* **aþwoge**.
      (p. 179)

      "Then he asked Oswy the king to send him some teachers, those who might
      convert his people to Christ's faith and **cleanse** them in the sanctifying wells
      of baptism."[6]

Although they look like singular verbs, *gecyrde* and *aþwoge* are actually plural, the subject *þa þe* referring back to the plural *lareowas*. Interestingly, even though B1 uses a late-West-Saxon spelling for *gecyrde* – the earliest manuscript, T, reads *gecerde* – he has preserved the denasalized reading in this position. The fact that denasalization is a relict feature is borne out by the distribution of this feature, as it

---

6.   All translations are my own.

most frequently occurs in the oldest manuscript, T, and in O, which retains several other conservative features.[7] Example (2), however, is unique to B:

(2) *Is þ sæd þ ða hæðenan ðritigu(m) siðu(m) mare werod **hæfde** þon(ne) osweo se cyning mid eahfriðe his suna.* (p. 186)
"It is said that the heathens **had** an army thirty times bigger than that of king Oswy and his son Ealhfrið."

Like many examples of denasalization found in other *Bede* manuscripts, this instance occurs in a subjunctive preterite plural, in a clause with reported speech. The subject *ða hæðenan* is a plural one, and the retention of a denasalized verb at this particular point is notable, given that B contains only a few other examples of denasalization. The fact that denasalization is often found in the same position in multiple manuscripts, and that it occurs most frequently in the earliest ones, suggests that it is a feature that was present in the original translation and that Example (2) is therefore a relict from the exemplar.

3.3 Double vowels

Grant is partly correct when he asserts that "there is, of course, no way to tell how much alteration of older or dialect forms had already been made in B's immediate exemplar" (1989: 13). However, B contains two scribal performances, and to an extent we can use them to control for some of the features we see in B. The differing treatment of double vowels by B1 and B2 is one example of this, and it suggests that at least some of these double vowels were present in B's exemplar.

As we can see from Table 1, in the part of Book 3 copied by B1 double vowels such as those in words like *tiid* and *riice* occur in all four *Bede* manuscripts, though they appear predominantly in T. Although O and Ca contain far fewer examples, these spellings nevertheless occur on occasion, with O preserving readings such as *uup* and the place-name *on briige*, which appears in all manuscripts. B's exemplar evidently also contained double vowels, because they have been copied by B1, and a selection of B1's examples is presented in the right-hand column.

Immediately noticeable is the distribution of double vowels in T and B. T often transmits multiple instances of a particular spelling: *tiid* occurs fifteen times in

---

7. In a number of cases, denasalized verbs occur in more than one manuscript; for example *onsende* (Miller 1890: 159: 9) occurs in T, O, and Ca; T and O share *meahte* (164: 9), *leornade* (224: 25), *worhte* (225: 25), *swulte* (250: 30), and *lifde* (252: 1). This feature is dealt with in more detail in Wallis (2013: 167–173).

Table 1. Double vowels in B1. The figures for T, O, and Ca are for the sections of the manuscripts corresponding with B1's copy. Figures in brackets indicate the total count per item for the whole of Book 3

	T	O	Ca	B1
tiid	9 (15)			1
too				1
uup(pan)		1 (1)		1
wool	5 (5)			1
on briige	3 (3)	1 (1)	1 (1)	3
in caale	1 (1)			1
diioma				1
good-	7 (7)			
riice	10 (15)			
liif	5 (7)			
hii	2 (2)	2 (2)	2 (2)	1

total, nine of these in the equivalent sections to B1's stint, and *wool* occurs five times. In contrast, items showing double vowels in B often occur only once, the exception being the place name *on briige*. A number of words which commonly have double vowels in T, such as *good-*, *rice*, and *liif*, do not appear with such spellings in Book 3 of B.

It is possible that B1 was more likely to retain a double vowel combination when it occurred in a proper noun, as over half of his examples belong to personal or place names, such as *on briige* (Faremoutiers-en-Brie, x3), *in caale* (Chelles, x1), and the personal name *diioma* (x1). This may have been motivated by a lack of familiarity with the names on B1's part; the place names are continental ones, and *diioma* is an Irish name. An interesting case is B1's treatment of the Scottish place name *hii* (Iona). This name appears three times in Book 3, and only in the part copied by B1. It seems that *hii* was not familiar to the later scribes of the *Bede*, as in one instance O and Ca (or the scribe of their exemplar) both write *his* in error (Miller 1890: 160: 2). *Hii* was evidently not a place name B1 knew well either; in the only place where he transmits it correctly (Miller 1890: 160: 2), the surrounding text overtly marks it as a name:

(3) *wæs he sended of ðam ealande 7 of ðam mynstre þe **hii** is nemned*   (p. 140)
     "He was sent from that island and from that monastery which is called **Iona**."

However, in two instances he fails to transmit the name faithfully. In the first, B1 erroneously substitutes *hibernia* for *hii* (4), while in the second, he omits the place name altogether (5):

(4) *æfter him fylgende wæs on ðone bysceophad finan se wæs eac fram **hibernia***
     scotta mynstre.   (pp. 164–165)
     "Finan succeeded him to the bishopric; he was also from **Ireland** (T: **Iona**), the
     monastery of the Scots."

(5) *7 hwearf eft on his eþel to **ðam mynstre** 7 ealande*   (p. 177)
     *hwearf eft on his eðel to **hii** þæm ealonde* (T: f.42r)
     "And returned to his homeland to **the monastery** and island (T: to the island
     of **Iona**)."

B1's motivation for the treatment of these two instances of *hii* is unclear, as *hibernia*
is clearly not the same as *hii*. While it is true that Iona was an Irish foundation and
B1 may have known this, the description of Ireland in Example (4) as *scotta mynstre*
"the monastery of the Scots" is undoubtedly clumsy, and the preservation of *hii*
would have made far better sense. It is possible that in Example (5) *hii* was omitted
in error, though none of the other manuscripts include the words *ðam mynstre*, and
it is tempting to speculate that they were substituted for the original name.

 If we are correct in assuming that these non-English personal and place names
were unfamiliar to B1, one strategy may have been to copy the names literatim. The
example of *hii* shows that he struggled to make sense of the reading, and the only
place where he transmitted it faithfully was where the text made it plain that it was
the name of a place. In the other two instances, the reading transmitted may be
due to an unsuccessful attempt to make sense of the exemplar text. Nevertheless,
B1 did retain a few double vowel spellings from ordinary nouns in his copy, such
as *tiid* and *wool*, though we will never know, of course, how many double-vowel
spellings existed in his exemplar.

 In contrast, B2 never writes double vowels in his section of Book 3. It is possible
that there were no such spellings in that part of the exemplar; however, T has 17
double vowel spellings in the section corresponding to B2's copy, and it therefore
appears that double vowels were a feature that B1 tolerated to a greater degree than
B2. The fact that double vowels were a feature of the original translation is suggested
by their presence in the earliest two manuscripts, T and Z. T contains 91 words with
double vowels in Book 3, while Z's three short excerpts from the *Bede* include *wiif*
(x2) as well as *wif* (x1) in its 214 graphic units. O and Ca preserve double vowel
spellings to a far lesser degree: O has 8 examples, while Ca has 5.

 As double vowels are not the only feature where there appears to be a difference
between B1 and B2 in scribal habit, this is certainly an area which will repay further
study:

> If a manuscript in more than one hand exhibits minor spelling variation between
> the hands, this points to idiosyncratic spellings being the work of the latest scribes,
> whereas a manuscript in a single hand which shows minor spelling variation be-
> tween its items suggests that its scribe has drawn items from several exemplars with
> different spelling practices.   (Scragg 1992: 351)

Scragg's comment also suggests that, in spite of Grant's (1989: 11) assertion that the differences between B1 and B2 were not worth distinguishing, there *is* a difference between the two in terms of their treatment of double vowels, which in turn shows the different reactions of the two scribes to their exemplar.

B1 is a scribe who, in spite of his tendencies to update the language of his exemplar to West-Saxon forms, does preserve occasional relict forms. In his habit of updating linguistic forms to late-West-Saxon ones, B1 is by far the most consistent of the *Bede* scribes in Book 3, and against this late-West-Saxon backdrop the few relict forms are notable.[8] The retention of some of these features, such as *f*-shaped <y>, must have been a conscious decision. The evidence of the double vowel spellings suggests that some of these forms were deliberately retained, especially when they were proper nouns belonging to unfamiliar people or places. Here the motivation for retaining relict forms may have been the lack of an alternative model for the spellings of these names. Nevertheless, B1 was not averse to making textual interventions and, in the case of *hii*, twice altered the text where the name appeared. Finally, some relicts may have been transmitted through oversight on B1's part. The ordinary nouns with double vowels and the denasalized verbs occur rarely, which suggests that despite their occasional presence in the exemplar, the scribe was not making a concerted effort to retain them.

4.   B1 as emender

In this section, we turn to some examples of B1's innovative scribal behavior. These examples show B1 emending his copy with varying levels of success, and they appear to show a scribe dealing with textual problems in the exemplar. In the first group of examples, the scribe makes successful alterations which result in a text that makes just as good sense as the original readings in other manuscripts, while in the second set word division appears to play a role in the difficult readings B1 presents. The final set of examples deals with emendations related to graphically based substitution, i.e. where the intervention appears to be based to some degree on the form and sequence of the graphs in the exemplar, and again we see that B1 sometimes struggled in his role as textual emender.

---

8.   B2's stint in Book 3 is not long enough to give a detailed picture of his scribal habits; nevertheless, we have been able to outline some broad differences between his approach and that of B1, such as their differing attitudes to double vowels and *f*-shaped <y>.

4.1   Successful interventions (lexical substitutions)

In the following examples, B provides a reading which differs from those given by the other *Bede* manuscripts. From this it is evident that B has been emended, as T, O, and Ca agree in their variants against B. However, we can view these alterations as successful textual interventions on the part of B1, as they provide readings which make good sense on their own, and without the evidence of the other manuscripts, it would be difficult to detect that a substitution had been made. In Example (6), B1 substitutes *wimmanna* "women" for the other manuscripts' *þinnenne* "female servants":

(6)   *⁊ sona þ geat ðæs mynstres ontynde. ⁊ eode mid anum hyre* **wimmanna** *to þara wæpnedmanna stowe.*   (p. 150)
   ⁊ hio sona þæt geat þæs mynstres ontynde ⁊ eode mid anre hyre **þinnenne** *to þære wæpnedmanna stowe.*   (O: f.37v)
   "And she immediately undid the gate and went with one of her **servant women** (B: **women**) to the men's building."

Although B's *wimmanna* is not as specific as *þinnenne*, it nevertheless makes good sense and fits both grammatically and semantically. It is possible that the change was motivated by confusion of <þ> and *wynn*, and subsequent reinterpretation of <nn> as <m>, although B1 is not generally prone to confusion of <þ> and *wynn*, and the alteration could be a simple word substitution.

   In (7), we also see a change of meaning:

(7)   *⁊ on missenlice wisan hit wann ⁊ wand ⁊ þræste ða sæminga becom hit on ða stowe þær se* **goda** *cyning ofslagen wæs.*   (p. 145)
   ⁊ on missen[…]lice dælas hit wond ⁊ þræste. þa semninga becom hit on þa stowe þær se **gemyngeda** *cyning ofslagen wæs.*   (O: f.35r)
   "And it struggled and writhed and twisted in different directions, then suddenly it came to the place where the **good** (O: **aforesaid**) king had been killed."

B's reading *goda* "good" for *gemyngeda* "aforementioned" again replaces the original word with a less specific one. Nevertheless, *goda* is an appropriate adjective to use for Oswald, whose miracle is being related at this point in the text, and it could be argued that the new reading is an improvement on the original. As with Example (6), the reason for the substitution is not obvious, though the replacement could have its roots in a textual misreading, perhaps through eyeskip and interpretation of the last four letters of *gemyngeda* (or a variant spelling such as *gemyn(d) goda*) as *goda*. This is certainly possible, especially if spacing had been irregular in the exemplar at this point or if the word had been split *gemyn(d)|goda* over two lines. The final example in this section is harder to explain away on grounds of mechanical error:

(8) *eall brytene cynn ⁊ mægða þe syndon on .IIII. **wereda** todælde þ is brytta. ⁊*
     peohta. ⁊ scotta. ⁊ angle   (p. 134)
     *all breotone cynn ⁊ mægðe þa syndon on feower **gereordo** todæled. þ is brytta ⁊*
     peohta ⁊ scotta. ⁊ angla   (O: f.29r)
     "All the people and tribes of Britain, who are divided among four **troops** (O:
     languages), that is Britons, Picts, Scots, and Angles."

B1's substitution of *wereda* "armies" for *gereordo* "languages" is more intelligent
than might appear at first sight. This passage narrates the expansion of Oswald's
kingdom to encompass speakers of four of the five languages Bede identifies as be-
ing spoken in Britain at the beginning of Book 1 (Miller 1890: 26). A word such as
wereda is not out of place semantically in a section dealing with military conquest,
and it may have appeared more appropriate to the scribe as he copied.

These three examples (along with Examples (12), (13), and, to an extent, (16)
discussed below) show that B1 was able to make sensible lexical substitutions which
were successful in so far as they retained the sense of the text, and would therefore
have been undetectable to a reader. In discussing similar substitutions by scribes
of poetic texts, Orton suggests that …

> The rejected word was [in these cases] not entirely opaque to the transmitter who
> altered or replaced it. This may be indicated when the substituted word has the
> same meaning as the one it seems to have replaced, showing that the transmitter
> might have been able at least to guess the meaning of the reading in his received
> text.   (Orton 2000: 99)

As we have seen, in two of the three examples above (*wimmanna* and *wereda*), B1
selects a replacement which maintains the meaning of the original text, even if it
does not occupy the precise semantic shade of the original. In the third case (*goda*),
the meaning is changed, yet as the new reading is appropriate in its context, we
cannot rule out the possibility that the change was intentional. The difficulty with
interpreting these examples is that we cannot know for certain what motivated B1
to make any of the alterations. Nevertheless, the text as it stands offers no prob-
lems of interpretation for the reader, and by that measure we can categorize these
interventions as successful.

4.2   Unsuccessful interventions (word division)

In the following examples, B1 emends the text unsuccessfully. By 'unsuccessfully',
I mean that he provides a text which gives an unsatisfactory reading because
the sense of the text is impaired. At first this might sit oddly with the notion
of a translator scribe who is keen to update spellings and some lexical items to

reflect his own preferred usage, a scribe who co-creates what Robinson describes as "the single most independent or revisionist version of the Old English *Bede*" (1981: 5). However, I argue that B1's occasional inability to produce a text which makes sense reveals something about the state of B's exemplar when its scribes copied from it.

In Example (9), a misunderstanding of the exemplar's word division seems to have been responsible for B's textual variation, with the possibility that dialect forms also occasionally contributed to B1's confusion. His textual alterations, while providing Old English words which exist, give us a text which in these examples fails to make sense:

(9)   *ða ongann heo on hyre mynstre cyrican timbrian **mare** eallre þara haligra apostola*
(p. 142)

*ongon heo on hire mynstre. cirican timbran. **in are**. ealra þara haligra apostola*
(T: f.31v)

"Then she began to build a church at her monastery, **in honor (greater than)** of all the holy saints."[9]

In Example (9), B1 has made two alterations to the text. Firstly, he seems to have mistaken the exemplar's *in are* "in honor" for the comparative *mare* "greater" (see Figure 4). A misreading of the minims is indicated by the fact that elsewhere in the manuscript he usually writes the West-Saxon variant *on* where other manuscripts have the non-West-Saxon *in*, and the first form appears to be his preferred usage. Secondly, he alters the genitive plural *ealra* to the dative singular *eallre*, interpreting the phrase as a dative of comparison (Mitchell 1985: Section 1360). This revealing course of action suggests that in this case, when faced with a problematic reading, his answer was to make the grammar of the surrounding text conform rather than to seek an alternative reading for *mare*. The reading in Example (9) indicates that in this particular instance, B1's copying and correction strategy was to look only locally for a solution to a textual problem and to assume that the answer lay in restoring grammatical concord.[10] This also suggests that B1 did not have access to (or chose not to consult) an alternative text of the *Bede* (either Latin or Old English) when making a choice about the reading in his exemplar.

---

9. "*Cum enim esset abbatisa, coepit facere in monasterio suo ecclesiam **in honorem** omnium apostolorum*" (Plummer 1896: 144).

10. For a parallel correction strategy, see Wallis (2016: 18–19).
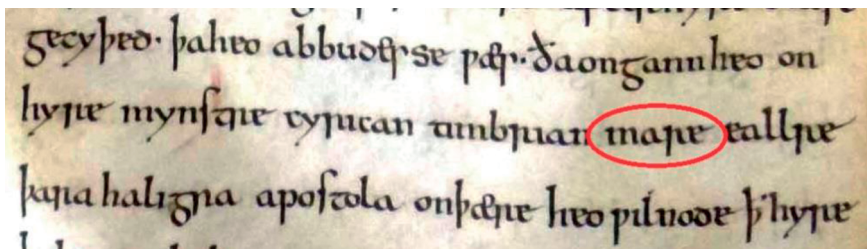
Figure 4. B1's substitution of *mare* for *in are* (p. 142)

A further instance of an unsatisfactory textual intervention occurs in (10):

(10) 7 *swa wæs geworden þæt se godes wer þurh witedomes gast þone storm **towearp næfre sæde**. 7 ðurh þæs ylcan gastes mægen þa upcumen wæs þet he hine aswefede 7 gestilde. þeah ðe he licumlice ðær **æfterwearð** wære.* (p. 162)
7 *swa wæs geworden þæt se godes wer þurh witedomes gast þone storm **toweardne foreseah**. 7 þurh þæs ylcan gastes mægen þa he uppcumende wæs 7 hine aswefede 7 gestillde þeah þe he licumlice þær **efweard** wære.* (O: f.43v)
"And so it happened that the man of god **foresaw the coming** storm through the spirit of prophecy. And through the strength of that same spirit when it was arising he soothed and stilled it, even though he was physically **absent** (was physically there **afterwards**)."[11]

Again, faulty word-division plays a part in the new reading. If B's exemplar had irregular spacing at this point, then this might account for B1's misdivision of *toweardne*, in reading *ne* as the beginning of the next word; Anglo-Saxon scribes do sometimes provide spaces between word-elements, and it would not be unusual to see a reading where a scribe had left a gap between the main word and a prefix or inflection (Orton 2000: 57–60). B1 may not have recognized or understood *foreseah*, as Campbell lists it as a specifically Anglian word (1951: 357). However, this fails to account for how the scribe dealt with the rest of the phrase. It is possible, given that several of the letters appear in the same place or sequence as in the other manuscripts (i.e. *næfre sæde* for *ne foreseah*), that the scribe was experiencing difficulty in reading parts of the exemplar at this point, and aimed to retain the letters he could make out with certainty. In addition, B replaces the word *efweard* with *æfterwearð*. A search of the *Dictionary of Old English corpus* for *æfterwearð/ æfterwearþ* yields no results. It therefore appears that B1 wrote <ð> in error for

---

11. "*sicque factum est, ut uir Dei et per prophetiae spiritum tempestatem **praedixerit futuram**, et per uirtutem eiusdem spiritus hanc exortam, quamuis corporaliter **absens**, sopiuerit*" (Plummer 1896: 158).

<d>.[12] This type of copying error is not unheard of in Anglo-Saxon manuscripts; Orton (2000: 23) notes an instance in *The death of Edgar* where the scribe wrote *weard* for *wearð*. If B1 did intend to write *æfterweard* for *efweard*, the substitution is puzzling, as *æfterweard* "afterwards" makes no sense in this context. It can perhaps be explained by B1's interpretation of *ef(t)* as an abbreviation for *æfter*, missing a suspension mark. This is not unfeasible, as the spelling *æft* occurs twice in B and once in T where other manuscripts read *eft*.

Example (11) is another instance of a problem prompted by irregular word-division.

(11) *7 forðon eanflæd seo cwen his mage for clænsunge his unrihtan sleges bæd osweo þone cyning. þæt he **for ge fæder** stowe mynster on to timrianne þam forespre-cenan godes ðeowe trumhere.*   (p. 188)

*7 forþon eanflæd seo cwen his mage fore geclænsunge his unrihtes sleges. bæd oswio þone cyning þæt he **þær forgefe** stowe mynster on to timbrienne þam foresprecenan godes þeowe trumhere.*   (O: f.55r)

"And therefore queen Eanflæd his kinswoman asked of King Oswiu that he should **donate** (*for ge fæder*) land there on which to have a monastery built by Trumhere, the aforementioned servant of God, as atonement for his unright-eous killing."

In the equivalent passage in O, the spelling of the verb *forgefe* lacks palatal diph-thongization, and it is possible, if B's exemplar also read *forgefe* rather than the West-Saxon *forgeafe*, that such a spelling may have triggered the misunderstanding. Although the other three manuscripts are unanimous in placing *þær* before *forgefe*, a position after the verb (*forgefe ðær*, for example) in the exemplar may then account for B's rather odd *for ge fæder*. The Latin confirms that *forgefe* must have been the original reading, because it translates *donaret*.

On a number of occasions, then, B1 provides a reading which comparison with the other manuscripts reveals to be an error due to the misdivision of words. It is possible that an exemplar with irregular spacing was the cause of such a scribal response. If this is the case, then it is notable that for these examples B1's skills only extended as far as making the subsequent letter strings into real Old English words and not as far as ensuring that sense was maintained at clause level.

---

12. As B1 also writes *stuðu* on p. 147 where the other manuscripts read *studu*, the appearance of <ð> for <d> above is not an isolated phenomenon in his scribal stint.

4.3 Graphically based substitutions

This final section considers a group of emendations which are unusual in that they seem to imitate the general shape of the word supplied in the other manuscripts. In several instances the new word makes just as good sense as the one it replaces. Orton (2000: 197) notes that in poetic manuscripts such substitutions happen often, and usually with an appropriate word chosen, as the meaning of the original is retained. Nevertheless, he concludes that "although many changes in this category suggest a fair understanding of the text, the motive behind them is often unguessable."

4.3.1 *Substitutions which retain meaning*

In Example (12), the only variation in the different manuscript readings is in the word's initial letter, and the meaning is not altered by the change. There is no dialectal difference between *leode* and *þeod*, and they both mean *people, a nation*. Their visual similarity may therefore have been one of the factors behind their interchangeability in B.

(12) *(W)ÆS þyses ylcan godes mannes gewuna þa he in east seaxu(m) bysceop þenunge brucende wæs. þ he gelomlice his agene **þeode** ⁊ mægðe norhimbraland sohte*

(p. 181)

*Wæs þysses yllcan godes monnes gewuna þa he on east seaxum bisceop þegnunge brucende wæs þæt gelomlice his agene **leode** norþanhymbra mægþe sohte*

(O: f.51r)

"It was the habit of this holy man, when he was undertaking his ministry among the East Saxons, that he frequently sought his own **people** (*leode/þeode*) and kin in Northumberland."

Likewise, (13) shows a similarity in the visual shape of the two phrases *æt gereorde* and *æt beode*:

(13) *ða dyde se broðor swa se oðer hine bæd ⁊ com eft on ham þa his gebroðro **æt gereorde** sæton.*   (p. 128)

*þa dyde þe broþor swa he hine bæd. ⁊ com eft on æfen ham. þa þa broþor **æt beode** sæton.*   (O: f.26r)

"Then the monk did as he was asked by his brother, and came back home to where his brethren sat **at their meal** (O: **at the table**)."

In this case, the substitution of *æt gereorde* "at [their] meal" for the other manuscripts' *æt beode* "at table" does not radically alter the meaning of the passage. While it is, of course, possible that B's scribe had elected independently to substitute one lexical item for another, it is interesting to note that the two words *gereorde* and

beode contain the same diphthong, end consonant, and (to an extent) rhyme. This raises the possibility that the substitution was made on an auditory level, i.e. that the scribe was writing from dictation, or that he was 'mishearing' the word in his head as he read from his exemplar and then copied his text:

> It may well be that in many such cases what happens is that the scribe moves from copying in a purely visual way to copying via 'the mind's ear'. Instead of reproducing a perhaps laboriously interpreted visual image, the visual image is now interpreted at a glance; and what is held in the mind between looking at the exemplar and writing down the next bit of text, is not the visual symbols, but the spoken words that correspond to them. What the scribe reproduces is then the words that he hears, not the visual images from which they arose: *regardless of whether his lips move, he is writing to his own dictation.*   (Benskin & Laing 1981: 6; emphasis mine)

The problem with viewing these examples as emendations at an audio level is that, although there are several cases where the sounds of the words may have been quite similar, these word pairs also have similar spellings or letter-sequences, and so a visual motivation for the change cannot be ruled out. From this point of view the emendations in this section bear resemblances to those discussed earlier, where alternative word division is combined with a similar sequence of letters in the new reading. As an alternative explanation, then, it is possible that B1 was copying from an exemplar which was not very easy to read and that he had to make a 'best guess' at its reading in some places. Whatever the reason for these substitutions, B's new readings in Examples (12) and (13) make sense, and like the successful lexical substitutions examined earlier, it is only by comparison with other *Bede* manuscripts or Bede's Latin that we become aware that a (successful) lexical substitution has taken place (Orton 2000: 47).

4.3.2   *Substitutions resulting in nonsense readings*
Despite B1's successful emendations in Examples (12) and (13), a number of other similar readings do not make sense. In these cases, B1 substituted for the original word one which fits badly with the sense of the text around it, yet which in some way imitates the shape of the word found in the other manuscripts. So, while in (14) below, *brædran* "broader" is obviously a poor choice for describing the *beteran* "most senior/respected" members of the monastery, it does share its initial letter and weak adjectival ending with the original:

(14) *þa ongann heo ymbgangan þa hus ðæs mynstres þara untrumra cristes þeowena*
7 swiðust þara ðe gelyfedre yldo oððe on gecorennesse heora þeawa maran 7
brædran *wæron.*   (p. 141)
þa ongan heo ymbgangan þa hus þæs mynstres. þara untrumra cristes þeowena.
7 swiðust þara þe gelyfedre yldo wæron. oððe on gecorenesse heora þeawa maran
7 **beteran** *wæron.*   (O: f.32v)

> "Then she began to go round the dwellings of the monastery of those servants of Christ who were sick, especially those who were advanced in age, or who were **greater and better** (**mightier and broader**) in the goodness of their conduct."

Similarly, *lare* and *lafe* are very similar in shape, and this substitution may have arisen through confusion of letter forms, between an <f> and <r> with long descenders:

(15) *Mid þy þe þa tyn dagas þæs feowertiglican fæstnes to* **lare** *wæron.*　　(p. 183)
　　　Mid þy þa tyn dagas þæs feowertilican fæstennes to **lafe** *wæron.*　　(O: f.52r)
　　　"When there were ten days of the forty-day fast still **left**(**to teach**)."[13]

In both cases it appears that (at least at some points in the manuscript) having a word which best matched the shape of that in the exemplar trumped notions of producing a text with good overall sense. Orton talks about similar variations in poetic manuscripts and says that "they are more than simple copying mistakes, because recognizable words (albeit in forms and positions quite inappropriate to the context) are put together from the wreckage of the original readings" (Orton 2000: 60).

The final example in this section is again an instance of rewording by B1, but this time it seems to have been triggered by the word *gleaunesse*, which occurs in T and Ca spelled with <ƿ> (*wynn*):

(16) *þa he sæmninga se man þ geseah* **þa ongan he mid scearpre geleafnesse on ðære stowe halignesse beon** *þær his hors swa hraðe gehæled wearð. 7 he þær tacen asette.*　　(p. 145)
　　　þa he þa se mon þ geseah. þa **ongeat he mid scearpre gleaunesse hwæthugu wundurlicre halignesse on þære stowe beon** *þær his hors swa hraðe gehæled wæs. 7 he þær tacen asette.*　　(O: f.35r)
　　　"When the man saw that, then **he perceived with keen wisdom what kind of wonderful holiness was in that place** (**he started to have keen belief in the holiness of that place**), where his horse was so quickly healed, and he set a token there."[14]

B1's mistaking *gleaunesse* "wisdom" for *geleafnesse* "belief" could have come about in one of two ways. The scribe could have mistaken <ƿ> for <f>, if his exemplar had forms of the graphs which were sufficiently similar. Alternatively, if the exemplar had a reading such as the one in O (*gleaunesse*), it is possible that he confused

---

13. "*Cumque X dies XL^{mae}* **restarent**" (Plummer 1896: 175).

14. "*Quo ille uiso, ut uir sagacis ingenii, intellexit aliquid mirae sanctitatis huic loco, quo equus est curatus, inesse; et posito ibi signo*" (Plummer 1896: 146).

<u> written for *wynn* for <u> written for /v/. Finally, it is not uncommon in the *Bede* manuscripts to see <ḡ> written for <ge>, leading to a misreading of *g(e)leaf-nesse/g(e)leaunesse* for *gleawnesse*. Additionally, B1 rewords the text, possibly to make his emendation make more sense (although *beon* sits rather awkwardly in the phrase). This suggests that his alteration was intentional, and that B1 went to some trouble to arrange the surrounding text to maintain some kind of sense. The question remains as to why the scribe would have made such an effort to change the text, unless he deemed it necessary.

Faced with a difficult reading in B, we are again left to speculate about the condition of the text from which it was copied; if B1 really is the reforming scribe suggested by his adoption of several late West-Saxon phonological, morphological, and orthographical features (Wallis 2013: 96–111), is it likely that unfamiliarity with a form alone would be the reason for the production of such a mangled reading? In some of the examples above (e.g. *mare* for *in are*, *geleafnesse*), B1 apparently tried to remedy a deficiency in the text caused by his new reading, which leaves us with two possibilities in considering the cause of these textual changes. Firstly, it is possible that B1 was an overzealous corrector of his text; however, if that were the case, we have to ask ourselves how satisfied a translator scribe would be with producing an incomprehensible text in these places. B1 certainly does not appear to be the kind of scribe to slavishly copy his exemplar, so it seems unlikely that he would choose to produce a nonsensical text. A second option is that B's exemplar was so poor that in attempting to remedy it, the scribe chose to follow as closely as possible what was written. As the text in front of him was illegible in places, he was unable to make good sense of his text and had to use all his resources – at times without success – to bring order to the text. If this was the case, then his attempts to make good the text (at least in some places) were unsuccessful.

5.   Conclusion

To conclude, then, it is clear that B1 is a translator scribe, in that he emends spellings on both a phonological and a morphological level. Nevertheless, he also retains a number of relict features, with varying levels of consistency. What is interesting about the conservative features he preserves is that they are so varied, as they include palaeographic relicts, as well as orthographical and morphological ones. For two of the features examined here, *f*-shaped <y> and double vowels, B1's usage differs from that of B2, which suggests that the distribution of the features is due to the latest scribes (B1 and B2), and not to their exemplar. It also demonstrates that the present scribes were responsible for at least some of the innovative features

associated with B, and which made it, in the eyes of many early scholars, a disappointing object of study as far as the Old English *Bede* was concerned. Furthermore, the presence of these relict features suggests that B's exemplar was at the rather conservative end of the *Bede* continuum in many ways, and perhaps an old manuscript by the time B was copied from it.

Additionally, there are instances where B1 acted as a textual emender, with varying levels of success. In some cases, he was able to make convincing lexical substitutions which are detectable only by comparing B with the text found in other manuscripts. In other cases, he was clearly less successful. In the places where B's variant reading makes little sense, problems seem to arise through a misdivision of words or through a partial reading which appears to retain the order and form of some of the graphs of the original. In these cases, we might speculate that such a problematic reading arose where the scribe had difficulty in construing the exemplar, for example where it had become illegible through age or damage of some kind. Therefore, although the scribe demonstrates a high level of innovation in some features and sometimes makes these substitutions well and fairly consistently, in other cases he transmits unsatisfactory readings based on the letters or spellings he could salvage from the exemplar. This course of action is supported by B1's action when dealing with some relict forms. In some of these cases (for example where there are double vowels), his transmission of relicts appears to be due to his unfamiliarity with some of the name forms he encounters, and the example of *hii* (and other proper nouns with double vowels) shows us that B1's tendency to interfere with and update his exemplar text was tempered by his retention of exemplar spellings when faced with unfamiliar names. This sheds valuable light on B1's behavior as an emender; faced with unfamiliar text, for which he was unable to provide a 'correct' reading according to his own training and scribal norms, he copied literatim. In grappling with the text and providing a reading which is clearly unsatisfactory, it appears that B1 resorted to transmitting as much of the text as he could by making the letters of the exemplar fit words he knew, even if the resulting reading made little sense. Far from being a "careless fellow" (Grant 1989: 10) and incompetent copyist, B1 reveals himself to be a scribe deeply engaged in updating and emending his text, using all his resources to copy from a challenging and possibly at times illegible exemplar.

References

Benskin, Michael & Margaret Laing. 1981. Translations and Mischsprachen in Middle English manuscripts. In Michael Benskin & Michael L. Samuels (eds.), *So meny people longages and tonges: Philological essays in Scots and Medieval English presented to Angus McIntosh*, 55–106. Edinburgh: Benskin & Samuels [published by the editors].

Budny, Mildred. 1997. *Insular, Anglo-Saxon and early Anglo-Norman manuscript art at Corpus Christi College, Cambridge: An illustrated catalogue*, 2 vols. Kalamazoo, MI: Medieval Institute Publications.

Campbell, Alistair. 1959. *Old English grammar*. Oxford: Clarendon.

Campbell, Jackson. J. 1951. The dialect vocabulary of the Old English Bede. *JEGP* 50. 349–372.

Dictionary of Old English corpus. http://quod.lib.umich.edu.eresources.shef.ac.uk/o/oec/index.html (31 October, 2015.)

Grant, Raymond J. S. 1989. *The B text of the Old English Bede: A linguistic commentary*. Amsterdam: Rodopi.

Ker, Neil R. 1957. *Catalogue of manuscripts containing Anglo-Saxon*. Reissued with supplement in 1990. Clarendon: Oxford.

Laing, Margaret. 2004. Multidimensionality: Time, space and stratigraphy in historical dialectology. In Marina Dossena & Roger Lass (eds.), *Methods and data in English Historical dialectology*, 49–93. Bern: Peter Lang.

Lapidge, Michael. 2008. The Latin exemplar of the OE Bede. In Patrizia Lendinara (ed.), … *un tuo serto di fiori in man recando: Scritti in onore di Maria Amalia D'Aronco*, 2 vols., 235–246. Udine: Forum 2008.

Miller, T. (ed.). 1890. *Bede's ecclesiastical history of the English people*. (Early English Text Society, Original Series, 95–96, 110–111). Oxford: Oxford University Press.

Mitchell, Bruce. 1985. *Old English syntax*, 2 vols. Oxford: Clarendon. doi:10.1093/acprof:oso/9780198119357.001.0001

Orton, Peter. 2000. *The transmission of Old English poetry*. Turnhout: Brepols.

Plummer, Charles. 1896. *Baedae opera historica*, vol. 1. Oxford: Clarendon.

Robinson, Fred. C. 1981. Bede's envoi to the Old English history: An experiment in editing. *Studies in Philology* 78(5). 4–19. (Reprinted in Robinson, Fred C. 1994. *The editing of Old English*, 167–179. Oxford: Blackwell.)

Rowley, Sharon M. 2004. Nostalgia and the rhetoric of lack: The missing exemplar for Corpus Christi College, Cambridge, Manuscript 41. In Joyce T. Lionarons (ed.), *Old English literature in its manuscript context*, 11–36. Morgantown: West Virginia University Press.

Rowley, Sharon M. 2011. *The Old English version of Bede's Historia Ecclesiastica*. Cambridge: D.S. Brewer.

Schipper, Jacob. 1899. *König Alfreds Übersetzung von Bedas Kirchengeschichte*. Leipzig: Georg H. Wigands Verlag.

Scragg, Donald G. 1992. Spelling variation in eleventh-century English. In Carola Hicks (ed.), *England in the eleventh century: Proceedings of the 1990 Harlaxton Symposium*, 347–354. Stamford: Paul Watkins.

Stokes, Peter A. 2014. *English vernacular minuscule from Æthelred to Cnut, circa 990–circa 1035*. Cambridge: D.S. Brewer.

Treharne, Elaine M. 2003. Producing a library in late Anglo-Saxon England: Exeter 1050–1072. *Review of English Studies* 54(214). 155–172. doi:10.1093/res/54.214.155

Treharne, Elaine M. 2007. Bishops and their texts in the later eleventh century: Worcester and Exeter. In Wendy Scase (ed.), *Essays in manuscript geography: Vernacular manuscripts of the English West Midlands from the conquest to the sixteenth century*, 13–29. Turnhout: Brepols. doi:10.1484/M.TCNE-EB.3.2761

Waite, Greg. 2014. Translation style, lexical systems, dialect vocabulary, and the manuscript transmission of the Old English Bede. *Medium Ævum* 83(1). 1–48.

Wallis, Christine. 2013. *The Old English Bede: Transmission and textual history in Anglo-Saxon manuscripts*. Sheffield: University of Sheffield dissertation. http://etheses.whiterose.ac.uk/id/eprint/5459

Wallis, Christine. 2016. Unpublished dry-point annotations in a manuscript of the Old English Bede: Oxford, Corpus Christi College 279B. *Medium Ævum* 85(1). 15–32.

CHAPTER 5

Old English wills
A genre study

Lilo Moessner

RWTH Aachen University, Germany

Following Bhatia's model of the 'language of the law', this paper treats wills and statutes as representatives of the genre 'written formal legislative legal documents'. In an overview of the publications on Present-day English (PDE) legal writing, I point out that only statutes have been analyzed with corpus-linguistic methods (Section 1). Quantitative data about Old English (OE) legal genres do not exist; OE wills have been investigated only from a sociohistorical point of view (Section 2). This research situation requires a new approach to the study of OE wills, which is quantitative-qualitative and corpus-based. The corpus consists of 23 wills, amounting to 10,330 words (Section 3). Its analysis is presented in two parts. In the first part, a new model of the text structure of OE wills is proposed; its constituents are described and illustrated by examples. In the second part, the linguistic features sentence length, sentence structure in terms of the number of clauses per sentence and of the type and position of subordinate clauses in complex sentences, pronominal reference to the testator, and the categories of tense, mood, modality, and voice of the verbal syntagm are analyzed. The quantitative results are compared to the corresponding figures of PDE statutes. This comparison yields two sets of linguistic properties; one set is genre-specific, the other is period-specific (Section 4). In Section 5, the results of the paper are summarized, and a future research agenda is derived from them.

Keywords: legal documents, OE wills, corpus, statutes, genre, text structure, linguistic features, genre-specific, period-specific

© 2018 John Benjamins Publishing Company

1.   Legal language

In his state of the art article on the 'language of the law', Bhatia (1987: 227) points out that …

> [t]he term 'language of the law' encompasses several usefully distinguishable genres depending upon the communicative purposes they tend to fulfil, the settings or contexts in which they are used, the communicative events or activities they are associated with, the social or professional relationship between the participants taking part in such activities or events, the background knowledge that such participants bring to the situation in which that particular event is embedded and a number of other factors.

In this study, Bhatia sets up a model for the classification of legal genres which comprises the parameters 'medium', 'setting', and 'style'. Along the parameter 'medium', he distinguishes between the values 'spoken' and 'written'; along the parameter 'setting', between the values 'academic', 'juridical', and 'legislative'; and along the parameter 'style', between the values 'frozen' and 'formal'. In this model, the genres wills and statutes figure as representatives of the same class. It is characterized in terms of written medium, legislative setting, and formal style.[1]

As legal documents may show a diversity of genres, law students must be familiarized with the writing styles of the different genres they will have to master in their professional careers. Genre analysis is therefore a prominent requirement in their training (Northcott 2013: 221).

The genre that has so far been described most thoroughly by linguists interested in modern legal English is the language of statutes (Gustafsson 1975; Bhatia 1993; Trosborg 1995; Kurzon 1997; Diani 2001; Hiltunen 2001; Williams 2005). Nearly all of these studies are based on large amounts of data and use corpus-linguistic methods. Among the genres which have attracted less attention are case reports (Bhatia 1993), contracts (Trosborg 1995), assignments (Danet 1985), jury instructions (Charrow et al. 1982), and wills (Finegan 1982). These studies are based on a small amount of data, and the features which are described as typical of the genres investigated are not supported by quantitative evidence.

For the lay person, it is the lexicon of legal documents which causes most difficulties. This explains why representatives of the legal profession pay much more attention to this aspect of their professional language than to its grammatical and discourse features. Mellinkoff (1963: 11) lists nine "chief characteristics of the language of the law", all of which concern lexical features, among them the use of

---

1.   Danet (1985: 276) proposed a similar but simpler model with just the parameters 'medium' and 'style'.

archaic words, loan-words from Latin and Old French, formal words, words with uncommon meanings, and words and expressions with flexible meanings. His book is not corpus-based, nor does it focus on one particular genre of legal documents; his examples come from law dictionaries, jury instructions, court cases, and several other sources.

2. Old English legal writing

Some genres of legal writing did not yet exist in OE (e.g. courtroom discussions, jury instructions); for others we do not have written records (e.g. witness examinations). The genres which have been described are law-codes (Hiltunen 1990; Schwyter 1996), documents containing lawsuits (Schwyter 1996), and wills (Danet & Bogoch 1992a, 1992b, 1994; Drout 2000; Tollerton 2011).[2]

Schwyter's (1996) comparison of the language of law-codes and lawsuits focuses on lexis, and in particular on the lexical field of theft. His approach here is quantitative, in that he provides figures for the frequency of theft-lexemes in texts representing each of these genres. With regard to grammatical features, however, his statements are observations rather than corpus-based results. Hiltunen's data on law-codes come from Liebermann's *Gesetze der Angelsachsen* (1903–1916). But the author does not provide details about the size of his corpus nor about the frequency of individual linguistic features, which he describes only vaguely as 'frequent' or 'rare'.

The majority of genre descriptions of OE legal writing concern wills. Yet these descriptions focus on sociohistorical issues, and when linguistic features are mentioned they serve to support the extralinguistic claims.[3] Thus, frequent reference to and care for the souls of the testators' ancestors in the afterlife is interpreted as a sign of the influence of the Benedictine Reform on will-making in the 10th century (Drout 2000). An address to the testator's lord with the request to allow the making of the will and/or to guarantee that its dispositions will be carried out after the testator's death is mentioned in support of the claim that Anglo-Saxon wills served to establish a close relationship between nobility and the royal court (Tollerton 2011). Binomial expressions like *mid mete and mid mannum* "with cattle and with men" and witness lists are interpreted as residues of orality, vague reference to

---

2. Schwyter (1996: 131) contrasts the different functions of law-codes and lawsuits as "prescription with general applicability versus description of individual legal cases."

3. Sheehan's (1963) occasional remarks on the use of particular expressions in OE wills (e.g. *sprece*, *cwide*) are not meant as contributions to our knowledge of the linguistic structure of OE wills. They are to be understood as landmarks in the history of the genre.

beneficiaries and bequeathed objects as well as dating with reference to contemporaneous events as indicators of context-dependency, and first-person reference to the testator as a sign of the incipient performative power of OE wills (Danet & Bogoch 1992a, 1992b, 1994). Danet & Bogoch (1992a: 102; 1994: 112) also present a quantitative distribution of first- and third-person reference to the testator.

3.   A new approach to research on OE wills

3.1   Method

The studies discussed in Section 2 have in common that they contain hardly any quantitative statements about characteristic linguistic features of OE wills. This research situation calls for a new approach to the study of OE wills. It is my aim in this paper to provide a comprehensive picture of the structural and linguistic properties of the genre of wills in OE, and I will adopt a corpus-based, quantitative-qualitative approach. The qualitative approach involves relating my analysis of the structural features to the findings of the sociohistorically oriented studies of OE wills. The quantitative analysis comprises those linguistic features for which quantitative results are available in studies on comparable legal genres in Present-day English (PDE). Here statutes, the equivalent of OE law-codes, are the appropriate genre. Like wills, they belong to the written medium, legislative setting, and formal style in Bhatia's classification model of legal genres.[4] Their linguistic features for which quantitative data are available include sentence length and syntactic structure, coordination and subordination of clauses, types and positions of subordinate clauses, person and function of pronouns, and the structure of verbal syntagms in terms of tense, mood, modality, and voice. The few quantitative statements about grammatical features of PDE wills (Finegan 1982) will be taken into account as well. Quantitative results about PDE legal genres will serve as points of reference, because quantitative studies of OE law-codes or other comparable legal genres do not (yet) exist.

3.2   Data

The standard edition of OE wills is by Dorothy Whitelock (1930). It contains 39 wills, most of which are also accessible on the internet in the *Electronic Sawyer: Online catalogue of Anglo-Saxon charters* (http://www.esawyer.org.uk). My data comprise the first 23 of Whitelock's wills. Those which are not downloadable from the internet (or not completely so) were keyed in manually. The corpus

---

4.   In Danet's model, the parameter 'style' has four levels, namely 'frozen', 'formal', 'consultative', and 'casual'. She classifies wills as frozen and statutes as formal (1985: 276).

thus established contains 10,330 words. It is neither annotated nor is its spelling standardized. Consequently the analysis had to be carried out manually. I will first describe the textual properties and then present a detailed analysis of the characteristic linguistic features of my data.

4. Corpus analysis

4.1 Text structure

The history of the vernacular will "from the time when the Anglo-Saxons first came under the influence of Christian missionaries until the second decade of the reign of Edward I" is traced by Sheehan (1963: 2), who mentions three constitutive parts: the notification, the disposition, and the sanction (1963: 59).[5] Since it is the purpose of a will to arrange what shall happen to the property of the testator after his/her death, the dispositions part is obligatory. The other two parts are characterized as optional but frequent. In the present paper, Sheehan's three-part model will be modified to fit my data better. The wills in my corpus contain two obligatory and up to three optional parts. The obligatory parts are the identification of the testators and the dispositions about their property; the optional parts are an address to the testator's lord, a list of witnesses, and a sanction.

4.1.1 *The testator identification*
All wills in my data start with the identification of the testator. It is contained in the first sentence, which follows one of these patterns:[6]

> *ic*, **x**, *wille biquethen / an / geswutelige / bidde …*
> "I, **x**, will bequeathe / grant / state / pray …"
> þis is **x's** *quide / gegurning*
> "this is **x's** will / desire"
> (*her*)**x** *cyð …*
> "(here) **x** makes known …"
> *her / on þis gewrite* **x** *geswutelaþ …*
> "here / on this document **x** states…"
> *her is geswutelod / on seo swuteling hu / þæt* **x** *…*
> "here [it] is declared / in the testimony / how / that **x** …"
> **x** *greteð* **y**
> "**x** greets **y**"

---

5. In the notification part, the person of the testator is identified.

6. Drout (2000: 12) lists similar patterns and calls them 'prefatory clauses'.

These testator identification patterns are illustrated in examples (1)–(6):[7]

(1) **Ic Wulfwaru bidde** minne leofan hlaford Æþelred kyning him to ælmyssan, þæt ic mote beon mines cwydes wyrðe.
"I, Wulfwaru, pray my dear lord King Ethelred, of his charity, that I may be entitled to make my will."

(2) **Þis is Ælfsiges biscopes cwide.**
"This is Bishop Ælfsige's will."

(3) **Her cyð \Æþeric/ on þissum gewrite** hwam he geann ofor his dæig þæra æhta þe him God alæned hæfð.
"Here in this document Æþelric makes known to whom after his day he grants the possessions which God has lent to him."

(4) **Her swutelað Wulfric** his hleofan hlaforde his cwyde 7 eallon his freondon.
"Wulfric here declares his will to his dear lord and to all his friends."

(5) **Her is on sio swutelung hu Ælfhelm** his are. 7 his æhta geuadod hæfð. for gode. 7 for wurulde.
"Herein is the declaration of how Ælfhelm has disposed of his property and his goods in fulfilment of his duties both to God and men."

(6) **Mantat ancer godes wræcca gretet Cnut king and Emma þe læfdie** swiþe bliþelike mid godes blisse.
"Mantat the Anchorite, God's exile, greets King Cnut and Queen Emma very joyfully with God's joy."

4.1.2   The dispositions

This part contains a list of items which are bequeathed; it may be arranged according to the beneficiaries or to the bequeathed objects. Among the beneficiaries are religious houses, the testator's lord, his family, and sometimes friends and/or servants. The list is introduced by variants of the formula *þæt is (þonne) ærest* "this is (then) first". For the action of granting, only a limited set of verbs is used, namely:

unnan	"grant, bestow, give"
becweðan	"bequeath, leave by will"
biddan	"ask, entreat, command, require"
willan	"will, wish, desire"

The objects of *becweðan* are expressed by a noun phrase, those of the verbs *biddan* and *willan* by a *that*-clause, and those of *unnan* by either a noun phrase or a *that*-clause.

---

7.   Translations not otherwise marked were adopted from Whitelock (1930).

The testators are not presented as disinterested parties, but they expect a countergift for their grants in the form of their commemoration and that of their families in church services, a burial at the expense of the religious houses which are the beneficiaries of their wills, or a dignified burial place. As pointed out by Tollerton, this feature, which testifies to a contractual or bilateral quality, is one of the differences from modern wills (2011: 53). In many wills, it is expressed explicitly, as in (7):

(7) *on ða gereðnesse is þæt land geseld to þam mynstre þæt man unc gefecce. æt uncrum ændedege mid þes mynstres crafte 7 unc swylce legerstowe forescewian swylc unc for gode þearflice sy. 7 for weorulde gerysenlic.*
"On these terms is the estate given to the minster: that on the day of our death they will fetch us with the minster's resources and provide for us such resting place as is necessary for us in God's sight and fitting in the eyes of the world."

Mostly the gifts are unconditional, but they may be restricted by previous gifts, or they may be valid only for the lifetime of the beneficiary or only when certain requirements are fulfilled, as exemplified in (8)–(10):

(8) *7 Ic geann minon fæder Æþelræde cynge. þæs landes æt Cealhtune buton þam ehta hidan þe Ic Ælmære minon cnihte geunnennen hæbbe.*
"And to my father, King Ethelred, I grant the estate at Chalton except the eight hides which I have granted to my servant Ælfmær."

(9) *on þet gerad þet he hæbbe þone bryce þes landes swa lange swa his tyma sy.*
"on condition that he is to have the use of the estate as long as his life lasts."

(10) *7 ic gean þæs landes. æt Brycandune. into sancte Petre. to Westmenstre. buton ic wylle þæt man mæste minum wiue. twa hund swyna. þænne þær mæsten sy.*
"and I grant the estate at Brickendon to St Peter's at Westminster, but it is my wish that when there is mast, two hundred pigs be fed for my wife's sake."

When a grant is restricted to the lifetime of a beneficiary, the will contains instructions as to what shall happen after the first beneficiary's death. The usual formula for the restriction of a grant to the beneficiary's lifetime is *his/hire dæge* "for his/her lifetime", and the next sentence begins with *æfter his/hire dæge* "after his/her death".

(11) *brucæ heo þæs landæs æt Batancumbæ* **hyræ dæg** *and* **æfter hire dæge** *ga hit an Ælfwærdes hand uncres suna*
"she is to possess the estate at Batcombe for her time and after her death it is to pass into the possession of our son Ælfweard"

The main grants in the wills are of a material nature, the regularly recurring term here being *land* "landed property". In addition, a great variety of objects are bequeathed

in the wills I studied, all in all nearly 100 items. They can be assigned to the lexical fields religious objects (*scrin* "shrine", *haligdom* "relics", *saltere* "psalter", *mæsse-reaf* "mass-vestments", *mæsse-boc* "mass-book"), war implements (*sweord* "sword", *spere* "spear", *handsæx* "a kind of axe", *helm* "helmet", *byrne* "coat of mail", *targa* "small shield, buckler", *wæpen* "weapon", *scield* "shield"), jewelry (*beag* "bracelet", *mentelpreon* "mantle-pin", *bula* "necklace", *hring* "ring", *bend* "band, ribbon"), animals (*steda* "stallion", *hors* "horse", *hafoc* "hawk", *heahdeorhund* "deer-hound", *hengest* "stallion", *oxa* "ox", *hriðer* "bull, heifer", *mearh* "steed", *colt* "colt"), clothes (*cyrtel* "coat, gown", *tunece* "tunic, coat", *mentel* "mantle, cloak", *crusene* "fur coat", *pæll* "silk robe", *fetel* "belt", *wifscrud* "woman's clothing", *hrægl* "dress"), and furniture and household gadgets (*myderce* "chest", *bedreaf* "bedclothes", *bedwahrift* "bedcurtain", *heallwahrift* "wall-tapestry", *sethrægl* "seat-covering", *hrægl-cyst* "trunk", *ruwa* "covering, tapestry", *andloman* "utensils, implements, vessels", *sopcuppe* "sopcup", *fæt* "vat, vessel, jar", *bledu* "dish-bowl, goblet", *hopscyte* "sheet, counterpane", *castenere* "cabinet, chest"). Among the unclassified objects are *wesend-horn* "bison's horn", *dryncehorn* "drinking-horn", *boc* "book, charter, title-deed", *scip* "ship", *seglgeræde* "sail-tackle", *blædhorn* "trumpet", and some others.

Apart from material goods, testators may give freedom to their bondsmen in their wills. They are named individually, or freedom is granted collectively.

(12) 7 *freoge man* **Wulfware** … 7 *freoge man* **Wulfflæde**
 "and Wulfwaru is to be freed … and Wulfflæd is to be freed"

(13) *man gefreogen* **ælcne witeþeowne man** *on ælcum þæra landæ þæ ic minon freondon bæcwedden hæbbæ*
 "every penally enslaved man be freed on each of the estates which I have bequeathed to my friends"

4.1.3   Address to the testator's lord

The address to the testator's lord is found in 13 out of the 23 wills in my data. It has no strictly fixed position in the wills, but preferred positions are at the beginning, right after the identification of the testator, or towards or at the end of the dispositions. The testator asks his lord for permission to make his will or asks him to guarantee that the will is implemented correctly. The usual formula here is *þæt min cwyde standan mote* "that my will may stand". Occasionally the testator asks his lord for the protection of his family or friends.

(14) *Þis is Æþelflæde cwyde þæt is ærest þæt ic gean minum hlaforde þes landes æt Lamburnan 7 þæs æt Ceolsige 7 æt Readingan. 7 feower beagas on twam hund mancys goldes. 7 .IIII. pellas. 7 .IIII. cuppan. 7 .IIII. bleda. 7 .IIII. hors. 7 ic bidde minne leouan hlaford for godes lufun.* **þæt min cwyde Standan mote***.*

"This is Æþelflæd's will. First, I grant to my lord the estate at Lambourn and those at Cholsey and at Reading, and four armlets of two hundred mancuses of gold, and four robes and four cups and four bowls and four horses, and I pray my dear lord for the love of God that my will may stand."

(15) *Nu bidde ic þone bisceop Ælfstan. **þæt he amundige mine lafe 7 þa þincg þe ic hyre læfe**. 7 gif him God lifes geunne lencg þonne unc **þæt he gefultumige þæt ælc þara þinga stande þe ic gecweden hæbbe***;

"Now I pray Bishop Ælfstan that he will protect my widow and the things which I leave her and, if God grant him longer life than us, that he will help to secure that each of the bequests which I have made may stand."

4.1.4   *The sanction*

This part, which appears at the end of a will, describes what may happen if somebody alters the will; it deals with the possibility that somebody may meddle with its text and with the punishments to be expected by the offender. Danet & Bogoch (1992b) refer to them as 'curses' and classify them along the parameters 'serious' vs. 'ludic' and 'conditional' vs. 'categorical'. Their corpus consists of 247 OE legal documents, of which 92 contain curses; 27 of these curses occur in wills, 44 in land grants and leases, and 21 in royal writs. My corpus contains 12 curses in 11 wills, and all of them belong to Danet & Bogoch's category 'serious *whoever* curses'. The first part, which specifies the offense (here the altering of the will), is expressed by a nominal relative clause or a conditional clause introduced by *gif*. The second part describes the punishment. In the wills in my corpus, the punishment takes effect in the next world. The offender must account for his offense before the Lord, and the consequences of the offense are described in more or less detail. They can be avoided if the offender makes amends before his death:

(16) *And wo so mine cuyde ofte*, *God him ofte heuene Riches*, *buten he it er his ende bete.*

"And whosoever detracts from my testament, may God deprive him of the kingdom of heaven, unless he make amends for it before his death."

(17) *And swa hwilc man swa þisne cwide awende*, *sy he Iudas gefere ðe urne drihten belewde on helle wite.*

"And whatsoever man shall alter this bequest, may he be a companion in the torment of hell of Judas who betrayed our Lord."

4.1.5   *The list of witnesses*

Tollerton (2011: 30–31) argues that witness lists "were particularly vulnerable in the copying process", and she assumes that this is the reason why few are extant in the versions of the wills available to us. In my corpus of 23 wills, only 4 contain lists of witnesses. Tollerton claims that their place is always at the end of a will, but I could

not make out a characteristic position – I found one at the beginning of a will, one in the middle of the dispositions, and two at the end of wills.

(18)   *7 þyses is to ywitnesse. Dunstan arcebisceop. 7 Aþelwold bisceop 7 Ælfstan bisceop.*
7 Æþelgar abbod. 7 se hired on Glestingabyrig. 7 þa twegen hiredas. on ealdan
mynstre. 7 on niwan mynstre on Wintanceastre.
"Of this the following are witnesses: Archbishop Dunstan, and Bishop Æthelwold, and Bishop Ælfstan, and Abbot Æthelgar, and the community at Glastonbury, and the two communities of the Old and New Minsters in Winchester."

4.2   Linguistic features

My selection of the linguistic features to be analyzed was guided, on the one hand, by what I observed to be recurrent patterns in my corpus and, on the other hand, by the list of features for which quantitative results are provided in studies on PDE statutes (see Section 3.1). This approach was chosen because quantitative studies of other OE legal genres do not yet exist, so that a synchronic genre comparison is not possible at present. Furthermore, the only publication on PDE wills (Finegan 1982) is based on four wills only and contains figures for only three linguistic features; a diachronic comparison between OE and PDE wills can therefore furnish only fragmentary results.

Where my results resemble those obtained for PDE statutes, there is a strong probability that the linguistic features concerned are typical of formal written legal documents produced in a legislative setting and that they did not change in the history of the English language. Where my results differ from those obtained for PDE statutes, I assume that my results are period-specific.[8]

4.2.1   *Sentence length*
I divided the texts in my corpus into sentences. This can be tricky, and several of my sentence divisions could be challenged. My two guiding principles were my own understanding of the text and Whitelock's (1930) understanding of the text as reflected in her PDE translations. Since the size of my wills varies from 32 to 1,003 words, it is not surprising that they also vary in their number of sentences, namely between 1 and 43. The overall number of sentences amounts to 428. Sentence length ranges from 4 to 106 words, with an average sentence length of 24.14 words.

---

8.   There is, of course, still the possibility that OE wills and OE law-codes, the forerunners of PDE statutes, differ with respect to their register properties.

In his pilot study of four PDE wills, Finegan (1982) reports an average sentence length of 39.8 words, as compared to an average sentence length of 19.3 words in the multi-genre *Standard corpus of present-day edited American English (= Brown corpus)*.

Sentence length in PDE statutes is discussed in Bhatia (1993), Gustafsson (1975), and Hiltunen (2001). Bhatia (1993: 106) mentions a law text comprising one single 271-word sentence and compares it to "a typical sentence in written scientific English" with an average of 27.6 words; this last figure is taken over from Barber (1988). Gustafsson (1975) analyzed the British *Courts Act* of 1971, which consists of 289 sentences with an average length of 55.11 words.[9] The shortest sentence is 10 words long, and the longest contains 179 words. In comparison, the genres 'scientific writing' and 'religion' in the *Brown corpus* have an average sentence length of 23.80 and 23.19 words respectively. Hiltunen (2001) analyzes a smaller corpus of five samples of five British Parliamentary Acts of the year 2000. Its 136 sentences have an average length of 45.05 words.[10] Comparing this figure with Gustafsson's, we note a decrease of the average sentence length by about 10 words per sentence over a period of 30 years. This is in line with what Kurzon reports; in his small corpus of three passages from three Finance Acts, the average sentence length drops from 92.50 in 1970 to 45.06 in 1980 and further to 37.06 in 1990 (Kurzon 1997: 131–132).[11]

The conclusion to be drawn from these figures is that the sentences of OE wills are shorter than those in the PDE legal genres wills and statutes. They are also shorter than the sentences of other PDE English for Specific Purposes domains, namely scientific writing and religious prose. Since the difference between the average sentence length of PDE wills and statutes is not very big, it is probable that this was so in OE as well. Under this hypothesis the small(er) average sentence length of OE wills is a period-specific feature of the written formal legislative genre.

4.2.2   *The number of clauses per sentence*

The number of clauses per sentence can be used as a measure of syntactic complexity (Hiltunen 2012: 43–44). The 428 sentences in my corpus consist of 921 clauses, which corresponds to an average of 2.15 clauses per sentence. Specifically, about one third of the sentences, namely 144, consist of one clause only, and only 17 (= 3.97%) contain 6 or more clauses. The most complex sentence, presented in (19), contains 8 clauses (the verbs of the clauses are underlined):

---

9.   The size of Gustafsson's corpus can be computed to contain nearly 16,000 words; she describes it as "42 book-size pages".

10.   Hiltunen's corpus contains a little more than 6,000 words.

11.   These figures are also quoted in Hiltunen (2012).

(19) *And ic <u>wille</u> þat men <u>nieme</u> þat erfe þat at Hoxne <u>stand</u>, þat ic þerto <u>bigeten</u>*
      <u>habbe</u>, and <u>dele</u> it man on to, half into þe minstre 7 <u>dele</u> for min soule. And <u>lete</u>
      men stonden so mikel so ic þeron <u>fond</u>, and <u>fre</u> man þo men alle for mine soule.
      "And I wish that one take the stock which is at Hoxne, which I have acquired
      there, and [that] one divide it into two parts, one half for the minster, and
      [that one] distribute [the other half] for my soul, and [that] one lets remain as
      much as I found there, and that one sets free all the men for my soul's sake."[12]

One sentence type, as in (20), does not meet our expectations about well-formedness;
it contains no main clause nor even a verb at all, and the missing verb cannot be
recovered from the linguistic context. It has to be supplied from the overall text
structure. Despite their deviant structure, sentences like (20) were counted among
the sentences containing one clause only.

(20) *Ærest his kynehlaforde ænne beah. on hundeahtotigan mancysan goldes; 7 an*
      handsecs. on ealswa miclan. 7 feower hors; twa geræded. 7 twa sweord gefetel-
      sode. 7 twegen hafocas; 7 ealle his headorhundas;
      "First, to his royal lord an armlet of eighty mancuses of gold and a short sword
      of the same value, and four horses, two with harness, and two swords with
      sheaths, and two hawks and all his staghounds."

This sentence is the beginning of the dispositions part in the will of Brihtric and
Ælfswith. It follows the identification of the testators and a list of witnesses. The
usual introductory pattern *ic an* "I grant" or *ic becweþe* "I leave" is missing.

My figures can be compared with those provided in Gustafsson's (1975) and
Hiltunen's (2001) studies on PDE statutes. Sentences in PDE statutes contain 2.86
clauses per sentence in Gustafsson's (1975) corpus and 2.52 clauses per sentence in
Hiltunen's (2001) corpus. Specifically, sentences consisting of one main clause only
take up a share of 20.76% in Gustafsson's (1975) corpus (60 sentences out of 289),
while Hiltunen (2001) reports a higher share of 25.7% (35 sentences out of 136).
Furthermore, sentences with 6 or more clauses amount to 6.92% in Gustafsson's
corpus and drop to 2.94% in Hiltunen's corpus. It appears from these figures that
sentence complexity (measured in terms of the average number of clauses per sen-
tences, the number of sentences consisting of one main clause only, and the number
of sentences with more than 6 clauses) decreases in statutes in the second half of
the 20th century, but that it is still higher than in OE wills. Like sentence length,
sentence complexity measured by the number of clauses per sentence turns out to
be a period-specific feature of the written formal legislative genre.

---

12. This is my translation, which tries to make the sentence structure transparent. Whitelock
(1930: 5) splits this sentence into two.

4.2.3   *Type and position of subordinate clauses*

Complex sentences contain several types of subordinate clauses, and these can be placed in initial, medial, or final position. In my corpus, I distinguished the following clause types: adjectival relative clause, noun clause, adverbial clause, and comparative clause; adverbial clauses are subdivided into conditional clauses, temporal clauses, purpose/result clauses, and causal clauses. Table 1 presents the distribution of the different types of clauses in OE wills.

Table 1. The distribution of subordinate clauses in my corpus of OE wills

Clause type	Absolute number	Percentage
Relative clause	208	41.86%
Noun clause	140	28.17%
Adverbial clause	94	18.91%
Conditional clause	56	
Temporal clause	19	
Purpose/result clause	18	
Causal clause	1	
Comparative clause	55	11.06%
Total	497	100.00%

Relative clauses have by far the largest share (41.86%), followed by noun clauses (28.17%). Adverbial clauses follow with 18.91%, and comparative clauses occupy the lowest rank on the frequency scale (11.06%).

The great majority of relative clauses are restrictive. They are used for the identification not only of pieces of land but of other bequeathed objects as well. The will of Ætheling Æthelstan is particularly rich in relative clauses; it contains 40 of the 208 relative clauses.

(21) *7 Ic geann minon fæder Æþelræde cynge. þæs landes æt Cealhtune buton þam ehta hidan þe Ic Ælmære minon cnihte geunnennen hæbbe. 7 þæs landes æt Norðtune. 7 þæs landes æt Mollintune. 7 þæs seolferhiltan swurdes þe Ulfcytel ahte. 7 þære byrnan. þe mid Morkære is. 7 þæs horses þe Þurbrand me geaf. 7 þæs hwitan horses þe Leofwine me geaf.*

"And to my father, King Ethelred, I grant the estate at Chalton except the eight hides which I have granted to my servant Ælfmær; and the estate at Norton; and the estate at Mollington; and the silver-hilted sword which belonged to Ulfketel; and the coat of mail which Morcar has; and the horse which Thurbrand gave to me; and the white horse which Leofwine gave to me."

Noun clauses usually function as objects governed by verbs like *an* "grant", *willen* "wish", *cyðen* "make known", *bidden* "pray, command", *þafien* "permit", etc. Sentence

complexity increases when several noun clauses cluster in one sentence as in (22). The first two instances of *þæt* introduce coordinated object clauses depending on *bidde*; the last instance of *þæt* introduces a lower-level object clause depending on *geþafige*.

(22) *Þonne bidde ic minnan leofan freond Ælfheah* *þæt* *bewite ægþer ge þa land ge þa þe mine magas sien, and* *þæt* *þu ne geþafige* *þæt* *man þis on ænig oþer wænde.* "Then I pray you, my dear friend Ælfheah, that [you] will watch both over the estates and those who are my kinsmen, and that you will never permit anyone to alter this in any way."

Whereas object clauses follow their matrix verb and adjectival relative clauses follow their antecedent, adverbial clauses have no fixed position in the sentence. They can occur in initial, medial, or final position. Table 2 summarizes the frequency and position of adverbial clauses in my corpus.

Table 2. Type and position of adverbial clauses in the corpus of OE wills

Clause type	Conjunction	Number of tokens	Initial	Medial	Final
Condition	*gif*	31	16	3	12
(*n* = 56)	*buton*	14		3	11
	on þe ræd þe	11		3	8
Time (*n* = 19)	*þe hwile þe*	11		5	6
	þa	2		2	
	þæt "until"	3		2	1
	þonne	2		1	1
	þær "when"	1		1	
Purpose/result	*þæt*	15		1	14
(*n* = 18)	*wiþ þon þe*	2			2
	buten "that not"	1			1
Cause (*n* = 1)	*for þon þe*	1			1
Total		94	16	21	57

Conditional clauses occupy the top rank of the frequency scale, and they are preferably introduced by *gif*. *Gif*-clauses occur with a frequency of 7.24 per 100 sentences.[13] All other types of adverbial clauses play only a minor role.

---

13. The ratio per 100 sentences was chosen to allow comparability with Finegan's frequency count for PDE wills.

Most adverbial clauses occur in final position. Initial position is occupied only by conditional clauses and only by those introduced by *gif*. When they occur in initial position, they are usually very short:

(23) *And gif he bern ne habbe; þanne go it into sancte Marie Stowe at Berkynge for vre aldre soule.*
"And if she has no child, then it is to go to St Mary's foundation at Barking for the souls of our ancestors."

The 11 instances of the complex conditional conjunction *on þe ræd þe* "on condition that" with its variants occur in medial but more often in final position.

(24) *And ic an þat Athelfled bruke þe lond þer wile þe hire lef beth <u>one raða</u> heo it on riht helde.*
"And I grant that Æthelflæd is to have the use of the estates as long as is agreeable to her, on condition that she holds it lawfully."

Conditional clauses introduced by *buton* express an exception of an action or state described before. Therefore *buton*-clauses are quite naturally excluded from sentence-initial position.

(25) *Gif hit þonne hwa do, God hine fordo ge mid sawle ge mid lichoman ge her ge on þam toweardan, <u>butan</u> ic hit self on oþer wænde.*
"If anyone do so, may God destroy him both soul and body, both here and in the future, unless I myself change it."

All comparative clauses are of the equivalence type (see Quirk et al. 1985: 1127). They are introduced by *swa* or *swylce*. As they require a standard of comparison, they are excluded from initial position. The 4 tokens of *swylce* occur in final position; of the 51 tokens of *swa*, 16 occur in medial position, 35 in final position.

(26) *7 ic gean Ælfhelme. 7 Wulfage þæra landa betwux Ribbel. 7 Mærse. 7 on Wirhalum. þæt heo hig dælan him betweonan. <u>swa hig efnost magon</u>. butan heora ægðer his agen habban wille. on þæt gerad þonne 'sceadd' genge sy. þæt heora ægðer sylle. iii. þusend sceadda. into þæra stowa æt Byrtune.*
"And I grant to Ælfhelm and Wulfheah the lands between the Ribble and the Mersey, and in Wirral, that they may share them between them as evenly as they can – unless either of them wishes to have his own – on condition that when it is the shad season, each of them shall pay three thousand shad into the monastery at Burton."

Sentence structure in terms of the types of subordinate clauses and their position in complex sentences is also described in Gustafsson's (1975) and Hiltunen's (2001) studies on PDE statutes. Apart from the share of comparative clauses, their findings

do not differ much from each other.[14] But they differ notably from the results obtained for my corpus of OE wills. OE wills contain a higher share of noun clauses than PDE statutes; conversely, PDE statutes have a higher share of relative clauses and adverbial clauses than OE wills. A comparison of the frequency of one particular class of adverbial clauses, namely conditional clauses, in OE and in PDE wills points into the same direction. Finegan's (1982: 117) study on PDE wills reports a ratio of 25.71 *if*-clauses per 100 sentences, the corresponding figure in my corpus is 7.24. The hypothesis to be derived from these comparisons is that the sentence structure of the written formal legislative genre changes with respect to the share of the different types of subordinate clauses involved. Specifically, the small number of adverbial clauses in OE wills is a period-specific genre feature.

Hiltunen (2001: 61–63) gives a detailed account of the positions of subordinate clauses in his corpus. His figures show a vast majority in final position (77.7%), whereas subordinate clauses in initial and medial position play only a minor role with 12.2% and 10.1%. Although final position is also preferred in my corpus, its 57 instances amount only to a share of 60.64%. Subordinate clauses in initial and medial position have a share of 17.02% and 22.34%. A comparison between the position of subordinate clauses in the OE wills of my corpus and in the PDE statutes of Hiltunen's corpus shows that the share of subordinate clauses in final position increased, whereas that in the other positions decreased. Like the type of subordinate clauses, their position is a period-specific feature of the written formal legislative genre.

4.2.4   *Pronominal testator reference*

In my set of OE wills, the prevailing testator reference strategy is by first-person pronouns. There are only four wills without a single occurrence of a first person (17.4%). On the other hand, first-person testator reference is the exclusive pattern in eleven wills (47.8%). In the remaining eight wills testator reference starts with third-person pronouns and then shifts to first-person pronouns (34.8%). Among these eight wills, three show just one occurrence of third-person testator reference at the beginning and then switch to first person. This result is in line with the figures provided by Danet & Bogoch (1992a: 102; 1994: 112) in their studies of 62 OE wills, and it confirms my implicit claim that my corpus is representative of the genre of OE wills.

In his study of PDE wills, Finegan (1982: 115) notes that "[f]irst person pronouns not only open each will but pepper them throughout." He points out that

---

14. Their figures are as follows: relative clauses (Gustafsson: 47%, Hiltunen: 50%), noun clauses (Gustafsson: 10%, Hiltunen: 14%), adverbial clauses (Gustafsson: 31%, Hiltunen: 34%), comparative clauses (Gustafsson: 14%, Hiltunen: 2%).

wills lack an addressee "because there is no way to know who will survive the testator." Finegan, therefore, finds it only natural that his data contain no instance of a second-person pronoun. However, since one of the optional parts of OE wills is the testator's address to his lord, the lord may be addressed directly and may be referred to not only by his proper name but also by a second-person pronoun. This is the case in six out of the thirteen wills that contain an address to the testator's lord.

In conclusion, we note that prevailing first-person testator reference is a feature which is shared by OE and PDE wills, but that the occurrence of second-person pronouns is characteristic of OE wills only.

4.2.5   *The verbal syntagm*

Finite verbal syntagms can be marked for one or several of the categories tense, mood, modality, and voice. I coded the 938 finite verbal syntagms in my data for each of these categories, and I will deal with them in this order.

With regard to tense, I distinguished in my analysis between verb forms referring to present time and those referring to a past time, i.e. simple past and present perfect. Modal (auxiliary) constructions are not coded for the category 'tense'; they will be treated separately. Table 3 shows the tense distribution in OE wills.

Table 3. Tense distribution in OE wills

Tense	Instances	Percentage
Present	715	79.80%
Simple past/Present perfect	181	20.20%
Total	896	100.00%

As can be seen, about four-fifths of all tense-coded verbal syntagms are present-tense forms. It should be no surprise that so many wills are written in the present tense, as they state what is to happen after the testator's death.

Williams (2005) describes the form of the verbal syntagm in PDE statutes in detail. His study is based on a corpus of 36 legislative texts from the major English-speaking countries, covering the first years of the present century and comprising 145,073 words. His corpus contains 7,168 verbal syntagms, of which 3,839 are coded for tense: 83.28% are present-tense forms, and 16.72% refer to past time. Williams interprets this distribution as a characteristic feature of prescriptive legal texts, because they "rarely look back into the past; they are essentially anchored in the present and projected towards the future" (2005: 76). He goes on to explain that in other legal genres, e.g. judicial decisions, past forms prevail.

In sum, the tense distribution in OE wills and in PDE statutes is so similar that it is safe to assume that present tense is not a period-specific, but a genre-specific feature of English written formal legislative documents.

With regard to mood, out of the 938 finite verb forms in my data, 731 are unambiguously marked for mood (indicative, subjunctive, and imperative).[15] The distribution of these forms is shown in Table 4.

Table 4. Mood distribution in OE wills

Mood	Instances	Percentage
Indicative	447	61.15%
Subjunctive	282	38.58%
Imperative	2	0.27%
Total	731	100%

The relatively large share of subjunctives may be surprising. They are unevenly distributed across the clause types, in that they occur most frequently in subordinate clauses, and particularly in conditional clauses and in clauses of purpose or result.[16]

In Williams' corpus, 3,887 verbal syntagms are marked for mood, and only 30 of them are subjunctive forms, which corresponds to a share of 0.77%.[17] Since it is a well-known fact that subjunctive frequency dropped in the history of English (Denison 1993: 330; Traugott 1972: 149) and that this development was reversed only recently (Övergaard 1995: 39), the relatively high frequency of subjunctives in OE wills has to be interpreted as a period-specific feature.

Modal (auxiliary) constructions in my corpus take up a share of 4.26% (40 tokens) of the verbal paradigms; all other verbal syntagms are realized inflectionally. The modal verbs involved are *willan* (15 tokens), *mot* (13 tokens), *sceal* (6 tokens), *mæg* (5 tokens), and *can* (1 token). In a legislative text one would expect a higher frequency of *sceal*, because this is the modal verb which expresses obligation (Traugott 1992: 195).

Williams (2005: 113) divides the finite verbal syntagms of his corpus into three classes: "(a) modals and semi-modals; (b) mood; and (c) indicative forms." The first group is very large with 3,281 tokens (45.77%); only indicative forms have a bigger share with 3,809 tokens (53.14%), and subjunctives and imperatives, which make up Williams' category 'mood', play only a minor role with 78 tokens (1.09%). The group of modal constructions is much more heterogeneous than that in the OE

---

15. In OE, the indicative : subjunctive contrast was not formally marked in the first person singular present tense. This explains the large number of ambiguous forms in my corpus.

16. The share of subjunctives in the 94 adverbial clauses is 73.40%, that in the 56 conditional clauses is 80.30%, and that in the 16 clauses of purpose or result is 88.89%.

17. The indicative is the prevailing mood with a share of 98.00%; the remaining 48 examples are imperatives (= 1.23%). All subjunctives in Williams' corpus are present tense forms.

wills; it contains not only the core modals *shall*, *should*, *may*, *might*, *will*, *would*, *can*, *could*, and *must*, but also the semi-modals *be to*, *ought to*, *be about to*, and *have to*. The most salient feature of this class is the high frequency of *shall*-constructions (1,621 tokens). They have a share of 49.4% of modal constructions and of 22.6% of all verbal syntagms. With reference to an earlier paper by Rissanen (2000), Williams reports that *shall* started its career as a characteristic feature of legal English in the 15th century. He argues that *shall* became so much of a hallmark of legal style that drafters of legal documents started to insert it even when no obligation was involved.

The low frequency of the modal *sceal* in OE wills for the expression of obligation is partly compensated for by the use of subjunctive mood, but the two constructions together cannot compete with the frequency of PDE *shall*-constructions. Like the relatively high frequency of subjunctives in OE wills, the low frequency of *sceal*-constructions is a period-specific feature.

All 938 verbal syntagms in my corpus are marked for the category 'voice'. Passive is represented by 22 verbal syntagms (2.35%); all others are in active voice. This is a surprising result, because passive constructions are usually considered typical of prescriptive legal texts. As Tiersma (1999: 76) points out, they serve to de-emphasize the actor.

With regard to PDE statutes, Williams (2005) does not provide any specific figures of passives. Yet he refers to an earlier paper of his, where he found that "one quarter of all verbal constructions in prescriptive legal English take the passive form" (Williams 2005: 35). For Finegan (1982: 118), passive marking is one of the outstanding features of PDE wills. His frequency count is, however, difficult to interpret: "66 passive verbs in an average 106 sentences per will."

Even if there are no precise quantitative statements available about the frequency of passives in PDE legislative texts, the low frequency of passives in OE wills has to be interpreted as a period-specific feature.

5.   Summary and conclusion

In the introductory sections of this paper, I introduced Bhatia's model for the classification of legal genres, and I surveyed studies of various PDE legal genres, pointing out that only statutes have been investigated with corpus-linguistic methods. With regard to studies of legal language in the OE period, besides law-codes and documents containing lawsuits, wills had been investigated in great detail. Yet the aim of the papers dealing with wills was to elucidate the sociohistorical importance of this genre.

In the present paper, I examined OE wills from a quantitative-qualitative and corpus-based perspective. It was my aim to provide a comprehensive description of OE wills in terms of their textual structure and linguistic features and to compare my findings to observations made in previous studies of OE and PDE wills and in studies of PDE statutes, the genre which in Bhatia's model is most closely related to wills. I established a corpus of 23 OE wills, comprising 10,330 words.

In the first part of my corpus analysis, I proposed a model of the text structure of OE wills with the obligatory parts testator identification and dispositions and the optional parts address to the testator's lord, sanction, and witness list. The *testator identification* in the first sentence of OE wills follows a fixed pattern. The *dispositions* specify the beneficiaries and the lands and/or personal objects to be bequeathed; the list of objects starts with variants of the formula *þæt is (þonne) ærest* "this is (then) first". The *address to the testator's lord* and *sanction* are found in my corpus in 13 and 11 of the 23 wills, respectively. *Witness lists* play only a minor part; they are present only in 4 wills.

In the second part of the corpus analysis, I presented quantitative data on sentence length, on sentence structure in terms of the number of clauses per sentence and of the type and position of subordinate clauses, on pronominal reference to the testator, and on the categories tense, mood, modality, and voice of the verbal syntagm. They were compared to the corresponding figures from studies of PDE statutes and – where possible – of PDE wills. It was assumed that similarities between OE wills and PDE statutes indicate genre-typical features of written formal legislative documents, whereas differences indicate period-specific features. On this hypothesis, the following features of OE wills were established as period-specific: shorter sentences, a lower average number of clauses per sentence, a higher number of sentences consisting of one main clause only, a higher share of noun clauses and a lower share of relative clauses and adverbial clauses, a less pronounced preference of adverbial clauses for final position, a larger number of subjunctives, and a smaller number of modal constructions and of passives. The prevailing use of present tense was identified as a genre-specific feature of legislative texts written in a formal style, and first-person testator reference as genre-specific of wills. Second-person testator reference proved a genre- and period-specific feature of OE wills.

These results are preliminary in as far as the genre of written formal legislative texts is represented for the PDE period mainly by statutes and for the OE period by wills only. They will be taken as the starting-point of a more comprehensive study of the genre of written legislative formal texts, which will include OE law-codes, the forerunners of PDE statutes, and PDE testaments, the successors of OE wills.

Corpus

Electronic Sawyer: Online catalogue of Anglo-Saxon charters. Devised and maintained by the Electronic Sawyer project team at the Center for Computing in the Humanities at King's College London. http://www.esawyer.org.uk (3 October, 2017).

References

Barber, Cesar L. 1988. Some measurable characteristics of modern scientific prose. In John Swales (ed.), *Episodes in ESP: A source and reference book on the development of English for science and technology*, 3–14. New York: Prentice Hall.

Bhatia, Vijay K. 1987. Language of the law. *Language Teaching* 20/4. 227–234. doi:10.1017/S026144480000464X

Bhatia, Vijay K. 1993. *Analysing genre: Language use in professional settings*. Harlow: Pearson Education Ltd.

Charrow, Veda, Jo A. Crandall & Robert Charrow. 1982. Characteristics and functions of legal language. In Richard Kittredge & John Lehrberger (eds.), *Sublanguage: Studies of language in restricted semantic domains*, 175–190. Berlin: Walter de Gruyter.

Danet, Brenda. 1985. Legal discourse. In Teun A. van Dijk (ed.), *Handbook of discourse analysis*, vol. 1, 273–291. London: Academic Press.

Danet, Brenda & Bryna Bogoch. 1992a. From oral ceremony to written documents: The transitional language of the Anglo-Saxon wills. *Language and Communication* 12. 95–127. doi:10.1016/0271-5309(92)90001-P

Danet, Brenda & Bryna Bogoch. 1992b. "Whoever alters this, may God turn His face from him on the day of judgment": Curses in Anglo-Saxon legal documents. *Journal of American Folklore* 105(416). 132–165. doi:10.2307/541083

Danet, Brenda & Bryna Bogoch. 1994. Orality, literacy, and performativity in Anglo-Saxon wills. In John Gibbons (ed.), *Language and the law*, 100–135. London: Longman.

Denison, David. 1993. *English historical syntax: Verbal constructions*. London: Longman.

Diani, Giuliana. 2001. Modality and speech acts in English Acts of Parliament. In Maurizio Gotti & Marina Dossena (eds.), *Modality in specialized discourse*, 175–191. Bern: Peter Lang.

Drout, Michael. 2000. Anglo-Saxon wills and the inheritance of tradition in the English Benedictine Reform. *Selim* 10. 3–53.

Finegan, Edward. 1982. Form and function in Testament language. In Robert J. Di Pietro (ed.), *Linguistics and the professions: Proceedings of the Second Annual Delaware Symposium on Language Studies*, 113–120. Norwood, NJ: Ablex Publishing.

Gustafsson, Marita. 1975. *Some syntactic properties of English law language*. Turku: University of Turku, Finland.

Hiltunen, Risto. 1990. *Chapters on legal English: Aspects past and present of the language of the law*. Helsinki: Suomalainen Tiedeakatemia.

Hiltunen, Risto. 2001. 'Some syntactic properties of English law language': Twenty-five years after Gustafsson (1975). In Risto Hiltunen, Keith Battarbee, Matti Peikola & Sanna-Kaisa Tanskanen (eds.), *English in zigs and zags: A festschrift for Marita Gustafsson*, 53–66. Turku: University of Turku, Finland.

Hiltunen, Risto. 2012. The grammar and structure of legal texts. In Peter M. Tiersma & Lawrence M. Solan (eds.), *The Oxford handbook of language and law*, 39–51. Oxford: Oxford University Press.

Kurzon, Dennis. 1997. 'Legal language': varieties, genres, registers, discourses. *International Journal of Applied Linguistics* 7. 119–139. doi:10.1111/j.1473-4192.1997.tb00111.x

Liebermann, Felix. 1903–1916. *Die Gesetze der Angelsachsen*. Halle: Max Niemeyer.

Mellinkoff, David. 1963. *The language of the law*. Boston: Little, Brown & Company.

Northcott, Jill. 2013. Legal English. In Brian Paltridge & Sue Starfield (eds.), *The handbook of English for Specific Purposes*, 213–226. Chichester: Wiley-Blackwell.

Övergaard, Gerd. 1995. *The mandative subjunctive in American and British English in the 20th century*. Stockholm: Almqvist & Wiksell.

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

Rissanen, Matti. 2000. Standardization and the language of early statutes. In Laura Wright (ed.), *The development of standard English 1300–1800*, 117–130. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511551758.008

Schwyter, Jürg R. 1996. *Old English legal language: The lexical field of theft*. Odense: Odense University Press. doi:10.1075/nss.15

Sheehan, Michael M. 1963. *The will in Medieval England*. Rome: Pontifical Institute of Mediaeval Studies.

Tiersma, Peter M. 1999. *Legal language*. Chicago: University of Chicago Press.

Tollerton, Linda. 2011. *Wills and will-making in Anglo-Saxon England*. Woodbridge: York Medieval Press/Boydell & Brewer Ltd.

Traugott, Elizabeth Closs. 1972. *The history of English syntax*. New York: Holt, Rinehart & Winston, Inc.

Traugott, Elizabeth Closs. 1992. Syntax. In Richard M. Hogg (ed.), *The Cambridge history of the English language*, vol. 1, 168–289. Cambridge: Cambridge University Press. doi:10.1017/CHOL9780521264747.005

Trosborg, Anna. 1995. Statutes and contracts: An analysis of legal speech acts in the English language of the law. *Journal of Pragmatics* 23. 31–53. doi:10.1016/0378-2166(94)00034-C

Whitelock, Dorothy (ed. & trans.) 1930. *Anglo-Saxon wills*. Cambridge: Cambridge University Press.

Williams, Christopher. 2005. *Tradition and change in legal English: Verbal constructions in prescriptive texts*. Bern: Peter Lang.

Spatio-temporal systems in Chaucer

Minako Nakayasu

Hamamatsu University School of Medicine

This paper presents a systematic analysis of synchronic spatio-temporal systems in Chaucer's language against the background of research in historical pragmatics and discourse analysis. The data consists of *The Canterbury tales* (fiction) and *A treatise on the astrolabe* (handbook). After defining spatio-temporal systems, I carry out a frequency study of the speaker's use of spatio-temporal elements, by comparing the two registers and contrasting proximal and distal elements. A qualitative analysis then highlights typical patterns of spatio-temporal elements in discourse. It reveals how proximal or distal elements combine in conveying a proximal or distal perspective and how these perspectives change in discourse.

Keywords: spatio-temporal system, Chaucer, historical pragmatics, discourse

1. Introduction

When a speaker decides to describe a particular situation in terms of language, he[1] must judge how far this situation is spatially and temporally removed from his deictic center (or his domain) and encode these spatial and temporal relationships with the aid of deictic elements. In the following fragment taken from Chaucer's Knight's tale, Theseus uses various expressions of space and time in explaining how the *Firste Moevere* "First Mover, or the God who created the world" created the world below him:

---

[1]   I refer to the speaker in general as *he*, not *she* or *they*, because the speaker in the upper level discourse structure is Chaucer himself (Pakkala-Weckström 2005; Nakayasu 2013).

(1)   *"The Firste Moevere of the cause above,*
       Whan <u>he</u> first <u>made</u> the faire cheyne of love,
       Greet <u>was</u> th'effect, and heigh <u>was</u> <u>his</u> entente ["plan"]. …
       <u>That</u> same Prince and <u>that</u> Moevere," <u>quod</u> <u>he</u>,
       "<u>Hath stablissed</u> in <u>this</u> wrecched world adoun ["below"]
       Certeyne ["definite, specific"] *dayes and duracioun*
       To al that <u>is</u> engendred in <u>this</u> place,
       Over the whiche day <u>they</u> <u>may</u> nat pace,
       Al ["although"] *<u>mowe</u>* ["can"] *<u>they</u> yet <u>tho</u> dayes wel <u>abregge</u>* ["shorten"] …"

                                                            (A.KN 2987–2999)[2]

Theseus (and the narrator) successively employ pronouns, demonstratives, tense forms, and modals,[3] considered 'deictic' elements (e.g. Bühler 1934; Fillmore 1997 [1975]) and constitutive of spatio-temporal systems.

The issues addressed here concern the nature of these spatio-temporal systems: what kind of expressions can be considered to be the elements of space and time, in what way they are connected to each other to form integrated spatio-temporal systems that convey a particular perspective, and how these connections change as the discourse progresses.

Although the idea of integrating spatial and temporal systems is not new, few studies have attempted an integrated analysis of spatio-temporal systems in historical data. Traugott (1974, 1978) was the first to employ the term 'spatio-temporal' in the analysis of historical data. Among the few studies available – Fries (1994) on text deixis in Early Modern English, Taavitsainen (1999) on personality and style of affect in historical data, and Nagucka (2000) on spatial and temporal meanings of *before* – the term 'spatio-temporal' is not employed. Most recently, I have examined spatio-temporal systems in Chaucer's *Treatise on the astrolabe* (Nakayasu 2015) and socio-pragmatic aspects of spatio-temporal systems in *The Canterbury tales* (Nakayasu 2014).

This paper carries out a systematic analysis of synchronic spatio-temporal systems in Chaucer's language, and of the elements making up those systems, along the lines of historical pragmatics and discourse analysis (Taavitsainen & Jucker 2010; Jucker & Taavitsainen 2013; Taavitsainen & Jucker 2015). The corpus used

---

2.   The following works were consulted on Chaucer's texts: Eisner (1975, 1976a, 1976b), Gunther (1977), Benson (1987), North (1988), and *MED* (2007). A.KN stands for 'The knight's tale' (Benson 1987); see Table 1 for abbreviations.

3.   Although scholars such as Bühler (1934) and Fillmore (1997 [1975]) do not include modals in their works on deixis, the present paper follows studies such as Halliday & Matthiessen (2014) and Langacker (2008) in incorporating modals in spatio-temporal systems (see Section 2).

is based on the Riverside edition (Benson 1987), consisting of selected tales from *The Canterbury tales* (34,102 words) and *A treatise on the astrolabe* (14,972 words), totaling 49,074 words, as is shown in Table 1:

Table 1. Chaucer's text making up the corpus used in the present study

Title	Abbreviation	Subtotal	Total
The Canterbury tales	CT		34,102
The general prologue	A.GP	6,686	
The knight's tale	A.KN	17,261	
The wife of Bath's prologue and tale	D.WB	10,155	
A treatise on the astrolabe	Ast		14,972
Total			49,074

In the *Helsinki corpus* (Kytö 1996), *The Canterbury tales* is categorized as 'fiction' and *A treatise on the astrolabe* as 'handbook' (astronomy). By including different types of texts representing the work of Chaucer, the present paper also compares the spatio-temporal systems in different registers in Chaucer. As Biber & Conrad point out, "effective register analyses are always comparative. It is virtually impossible to know what is distinctive about a particular register without comparing it to other registers" (2009: 36).[4]

The structure of each text should also be kept in mind when analyzing the systems (see also Nakayasu 2013: 49–50; 2015: 244–245). *The Canterbury tales* consists of four levels of discourse (Pakkala-Weckström 2005: 46–48): Level 1 (author [Chaucer] and reader), Level 2 (narrator [Chaucer the pilgrim] and narratee), Level 3 (the pilgrims),[5] and Level 4 (the characters in the pilgrim's tales). In *A treatise on the astrolabe*, on the other hand, there are only two levels: the first level is the author (Chaucer himself), and the second level is the narrator (Chaucer) addressing his ten-year-old son.

The discussion starts with a definition of spatio-temporal systems and the elements which constitute the system, incorporating the idea of deixis (Nakayasu 2014, 2015). Section 3 presents a quantitative analysis of how frequently these elements are employed by comparing fiction and handbook and contrasting proximal and distal elements. In Section 4, a qualitative analysis of the systems in discourse is

---

4.   Gather (this volume) and Suhr (this volume) reveal features of a genre, i.e. congregational songs and sensationalist news pamphlets, respectively, by comparing them with those of other genres. Moessner (this volume), in contrast, highlights features (genre- and period-specific) of OE legal writing by comparing them with PDE.

5.   Or the second narrators, following Nakayasu (2013).

provided discussing how these elements are interrelated with each other in taking a proximal or distal perspective and how these perspectives change in discourse. Finally, Section 5 summarizes the main points and advance implications for future research.

2.   What are spatio-temporal systems?

First of all, spatio-temporal systems are deictic in nature (Bühler 1934; Fillmore 1997 [1975]; Levinson 1983; Martin 1992; Duchan et al. 1995; Brisard 2002; Martin 2015). From numerous entities and situations in our world, the speaker selects some to encode in language, thereby assessing how far these entities and situations are from his domain, or deictic center. The entities close to this deictic center are referred to as 'proximal', while those distant from it are 'distal' (e.g. Diessel 1999). In addition to prototypical distance (i.e. spatial and temporal distance), distal elements can express metaphorical and hypothetical distance (Oakeshott-Taylor 1984; Nakayasu 2009).

Second, spatio-temporal systems encompass a variety of linguistic elements or expressions that refer to space and time; they are listed in Table 2:

Table 2. A survey of expressions of space and time

Category	Proximal	Distal
Pronoun	1st person (*I, we*)	3rd person (*he, she, it, they*)
	medial: 2nd person (*thou, ye*)	
Demonstrative	*this, these*	*that, those*
Adverb (spatial)	*here*	*there*
Interjection	*allas, O*	
Tense	present/non-past	past
Modal	*shall*	*should*
	will	*would*
	can	*could*
	may	*might*
	must (proximal)	*must* (distal)
Adverb (temporal)	*now*	*then*

Spatio-temporal elements which typically express spatial relations are pronouns, demonstratives, and adverbs.[6] Whereas the latter two show a dichotomy of proximal and distal, the former assumes a trichotomy of the first (proximal: speaker),

---

6.   This paper focuses on the contrast between *here* and *there* and excludes other spatial adverbs such as *yonder* from the analysis. Likewise, the analysis of temporal adverbs highlights the contrast between *now* and *then*, excluding other adverbs such as *tomorrow*.

second (medial: hearer), and third persons (distal: other).[7] The second-person pronouns play a crucial role in the speech situations (Halliday & Hasan 1976). The speaker chooses either the *th*-form (*thou*) or *y*-form (*ye*) according to the inter-actional status[8] of the speaker and the hearer (Jucker 2006). Interjections are also part of spatio-temporal systems;[9] they are neither precisely spatial nor temporal but rather constitute an integrated spatio-temporal category. In addition, they can only be proximal, because they are directly connected to the speaker's deictic center.

 Temporal relations are expressed most typically by the tense system. The present (non-past) tense is proximal, while the past tense is distal, representing distance from the speaker's domain. Unlike scholars such as Bühler (1934) and Fillmore (1997 [1975]), I argue here that modals, like tense forms, also belong to deictic, spatio-temporal systems.[10] In a way similar to tense, modals are frequently employed in utterances to make 'the proposition finite' (Halliday & Matthiesson 2014: 144), i.e. to relate the proposition (or 'situation' in the present paper) to the speaker's domain (or deictic center).[11] Halliday & Matthiesson (2014: 144) explain this significant function of finite elements such as tense and modality (or modals) in terms of the notion 'interpersonal deixis' because "they locate the exchange within the semantic space that is opened up between speaker and listener." Fritz takes a semiotic and speaker-oriented view that the system of modality (and also the future) is "based on a speaker pointing deictically to himself" (2003: 135). Both views emphasize that when a speaker employs a modal, he locates a situation on a time axis, judging how distant it is from his speaker domain. As this is a premise of spatio-temporal systems, modals are, in this particular sense, qualified to be part of spatio-temporal systems.[12]

---

7.   In addition to strongly deictic uses (i.e. pointing directly to the [extralinguistic] referent), third-person pronouns and demonstratives also show many anaphoric uses, i.e. referring to the previously mentioned entity in the context. In the present paper, whose purpose is to analyze overall spatio-temporal systems, these strongly deictic and anaphoric uses are included in the same category for simplicity.

8.   The interactional status of interactants consists of three aspects: social status, relation of the interactants, and situational status (Jucker 2006: 62).

9.   For the deictic nature of interjections, see Wilkins (1995).

10.   Moessner (this volume) treats modals separately from tense "because past forms of modal auxiliaries do not always refer to past time."

11.   In cognitive linguistics, this phenomenon is known as 'grounding' (Brisard 2002). Tense and modals are employed to situate "the profiled relationship with respect to the speaker's current conceptual reality" (Langacker 2008: 259).

12.   The meanings and functions of modals are complex in that, in addition to their original lexical meanings, modals involve the overlapping domain of modality and tense and in that they underwent historical change (Traugott 1972; Coates 1983; Palmer 2001; Nakayasu 2009). The

Third, one should not consider spatio-temporal systems as a mere accumulation of spatial and of temporal systems; rather, I argue that these systems are interrelated such that they constitute an integrated system. To that effect, expressions from spatial and temporal systems may go hand in hand, both when the speaker takes a proximal perspective and when he takes a distal perspective. If the speaker takes a proximal perspective, proximal elements of both space and time may systematically appear in the same context; similarly, when he takes a distal perspective, distal elements denoting space will coexist with elements denoting time. Moreover, he may also alternate these proximal and distal perspectives with each other. On this approach, then, the complexity of multiple dimensions of space and time is reduced to an integrated spatio-temporal domain.

From this definition of spatio-temporal systems, we can expect a number of benefits. First and foremost, by disregarding the overly elaborate details of each spatio-temporal element, such as the difference between the original meaning and metaphorically extended meanings in the temporal domain as in the cases of distal modals, new insights into the principle of regulating spatio-temporal systems may be obtained. Next, an analysis of integrated spatio-temporal systems may reveal which perspective, proximal or distal, the speaker is likely to take in a particular context, irrespective of whether he focuses on spatial or temporal systems or on the integrated system. In addition, the analysis shows the dynamics of change in discourse. Furthermore, it will become easier to contrast systems across different registers (Biber & Conrad 2009) and to follow the development of systems in history. Finally, this approach to integrating spatial and temporal systems allows a more systematic examination of meaning extensions from the spatial to the temporal domain, which are in fact attested quite frequently (Nagucka 2000, among others). To examine these points, particularly (i) to analyze the perspectives likely to be taken and the changes in discourse viewed from those perspectives and (ii) to compare different registers, the present paper presents quantitative and qualitative analyses of the elements which constitute these systems.

---

temporal reference of proximal modals is non-past, i.e. present or future. Distal modals may signify prototypical distance (temporal distance, i.e. past time reference), metaphorical distance (psychological distance), and hypothetical distance (distance from reality) (Oakeschott-Taylor 1984; Nakayasu 2009). The time reference of distal modals is also non-past in many cases where the speaker favors metaphorical and hypothetical distance, not temporal distance. In the present paper, which aims at a systematic analysis of overall spatio-temporal systems, I opted for an analysis that does not go deeply into the complexity of the various meanings and functions of the modals.

3.   Elements of space and time

3.1   Frequency analysis

Table 3 presents the frequencies of elements belonging to the spatio-temporal systems employed in the corpus, thereby comparing the registers 'fiction' (CT) and 'handbook' (Ast). Normalized frequencies (per 10,000 words) are given for each element.

   Of all the pronouns, the third-person pronouns are employed most frequently, due to the notably high frequency in CT. Indeed, third-person pronouns are exploited recurrently by the narrators to refer to the characters who participate in the tales (see Section 1 for the levels of discourse). First-person pronouns are used more often in CT than in Ast because the narrators refer to themselves and directly quote the characters' speeches. It should also be noted that second-person pronouns are used more often in Ast and that all of them are *th*-forms; this is because the speaker is addressing his 10-year-old son, explaining his astrolabe. Demonstratives appear more often in Ast, whose role as a scientific handbook may have motivated the speaker to specify which entity he is referring to.[13] It is also interesting to note that Ast employs the demonstrative *thilke* "this/these/that/those" more often, which does not make a distinction between proximal and distal reference. Regarding spatial adverbs, the distal adverb *there* is employed more often in Ast, because the speaker wants to specify which location he is referring to, somewhere in the hearer's domain in particular. In CT, in contrast, the narrator often introduces a new participant in his story with the aid of existential *there*. Interjections appear only in CT. The register 'fiction' can be considered to have an affinity with these interjections, which are directly linked to the speaker's domain, while the register 'handbook' cannot be.[14]

   Turning to the spatio-temporal elements of time, temporal relations are primarily represented by tense forms, in particular the simple present and the simple past. In CT, story-telling naturally encourages the narrators to make use of the simple past frequently as well as of variety of tense forms that refer to past time, such as the present perfect and the past perfect. The simple present is here often employed for specific purposes, such as to represent direct speech dialogues between characters (see the levels of discourse in CT; Section 1) or, as the so-called historical present, to

---

13. In Ast, the speaker provides explanations, sometimes pointing to illustrations, which may lead to the use of demonstratives, particularly proximals.

14. Although a study of the historical development of demonstratives is beyond this paper, Biber et al. (1999: 274–275) found that *this/these* are more frequent in academic prose than in fiction in Present-day English, while for *that/those* it is the other way around.

Table 3. Frequency of the spatio-temporal elements

	Proximal			Medial/Other			Distal		
		CT	Ast		CT	Ast		CT	Ast
Pronoun (*n* = 5235)	1st person	371.2	146.9	2nd person	148.4	234.4	3rd person	758.9	203.0
Demonstrative (*n* = 731)	*this/these*	95.3	112.9	*thilke*	6.2	28.1	*that/those*	26.7	55.4
Spatial adverb (*n* = 258)	*here*	7.3	2.0	Existential *there*	28.9	6.0	*there*	18.7	39.4
Interjection (*n* = 54)	(various)	15.8	0.0	–	–	–	–	–	–
Tense form (*n* = 4314)	Simple present	407.6	557.0	Present perfect	39.3	9.4	Simple past	505.2	85.5
				Other	0.6	0.0	Past perfect	26.1	0.0
	Total	407.6	557.0	Total	39.9	9.4	Total	531.3	85.5
Modal (*n* = 869)	*shall*	35.2	29.4	Other	5.6	0.0	*should*	16.1	2.0
	will	41.1	20.0				*would*	29.9	4.7
	can	15.0	0.7				*could*	15.5	0.7
	may	34.6	24.7				*might*	9.4	1.3
	must (proximal)	10.3	2.7				*must* (distal)	4.4	0.0
	Total	136.1	77.5	Total	5.6	0.0	Total	75.7	8.7
Temporal adverb (*n* = 336)	*now*	29.3	5.3		–	–	*then*	11.1	112.9
							tho	1.2	11.4
	Total	29.3	5.3	–	–	–	Total	12.3	124.2

refer to past events (Benson 1961; Nakayasu 2013).[15] In Ast, in contrast, the speaker strongly prefers the simple present which he uses to explain facts and to show his son how to use the astrolabe; therefore the variety of tense forms is quite limited. Modals appear generally more frequently in CT than in Ast.[16] In CT, the narrators extensively exploit a greater variety of modals to make their tales more vivid.[17] In particular, they employ the distal modals *should*, *would*, and *could* much more frequently than in Ast. The speaker in Ast, on the other hand, mainly uses *shall*, *will*, and *may*, because the scientific and dialogic characteristics of this handbook direct his attention to his own domain and scientific facts and relationships. The temporal adverbs *now* and *then* exhibit an interesting contrast between the two registers. In CT, *now* is employed more often to involve the hearers into the story, whereas in Ast, *then* is used much more frequently to enable the speaker to give step-by-step, logical explanations. In many cases, and this applies to *then* in particular, the meaning and function is metaphorically extended from typical temporal meanings. The adverb *tho* "at that time", on the other hand, carries a temporal meaning. This distinction between *then* and *tho* is in line with the definitions in the *Middle English dictionary* (*MED* 2007).

3.2 Proximal vs. distal elements

Now that the elements of space and time have been identified and their frequency distribution has been presented, the next step is to examine which perspective, proximal or distal, is taken in the corpus, and in which contexts. In Figures 1–4, the spatio-temporal elements given in Table 2 are classified in terms of the contrasting categories 'proximal' and 'distal'.[18] For the sake of convenience, a positive percentage share is assigned to proximal elements and a negative percentage share to distal elements. Figure 1 indicates the relative share of pronouns employed in the corpus in terms of distance from the speaker's domain, i.e. proximal or distal:

---

15. In Table 3, the historical present is counted as simple present.

16. The present paper covers all occurrences of modals, including 13 main verb uses. I aim to capture the full range of factors which the modal could cover, regardless of whether they function as main verbs or auxiliaries. This reasoning also underlies the decision to include both tense forms and modals in spatio-temporal systems. For more on modals, see Nakayasu (2009).

17. Other studies revealed that in other periods as well, modals are used more frequently in drama/fiction than in academic prose; see Biber (2004); Biber et al. (1999: 486–487); Biber & Conrad (2009: 94, 170).

18. The elements which are not distinctively either proximal or distal (for example, *thilke* and the present perfect) are excluded from the figures.
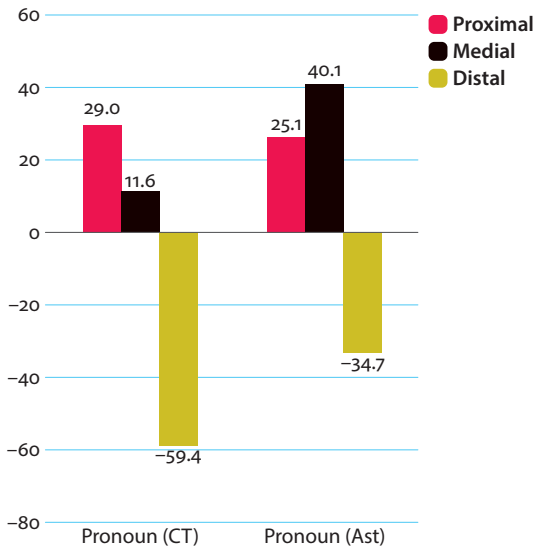
Figure 1. Proximal and distal pronouns

As seen in Section 2, pronouns involve a trichotomy of proximal, medial, and distal pronouns. As can be seen from Figure 1, medial pronouns have been included in the proximal domain.[19] Figure 1 furthermore shows that the percentage shares of proximal pronouns *sensu stricto* are similar in CT and Ast but that medial pronouns are used more frequently in Ast than in CT. As a result, proximal pronouns *sensu lato* (i.e. proximal and medial pronouns) take up a higher share than distal pronouns in Ast than in CT (65.3% vs. 40.6%). These figures tie in with the fact that the main purpose of the speaker in Ast is to explain the astrolabe to his son, which contributes to the high frequency of medial pronouns, while in CT, the narrators talk about the characters who participate in their tales often employing distal pronouns (see Section 3.1).

The ratios of proximal and distal demonstratives and spatial adverbs are presented in Figure 2.

In both CT and Ast, proximal demonstratives (*this/these*) take up a higher share than the distal ones (*that/those*), the ratio in CT being slightly higher. Conversely, distal adverbs (*there*) outnumber proximal ones (*here*).[20] The distal adverbs in Ast,

---

19. This is in line with Halliday & Hasan (1976), who argue that the person system derives its significance from the person's (or object's) relevance to the speech situation. That is, they distinguish the roles of speaker and addressee from other roles than speaker and addressee. Huang argues that in a three-term system, i.e. proximal, medial, and distal, "the middle term specifies a location also relative to the deictic centre" (2014: 195–197); see also Diessel (1999: 39).

20. Here, 'existential *there*' is excluded.
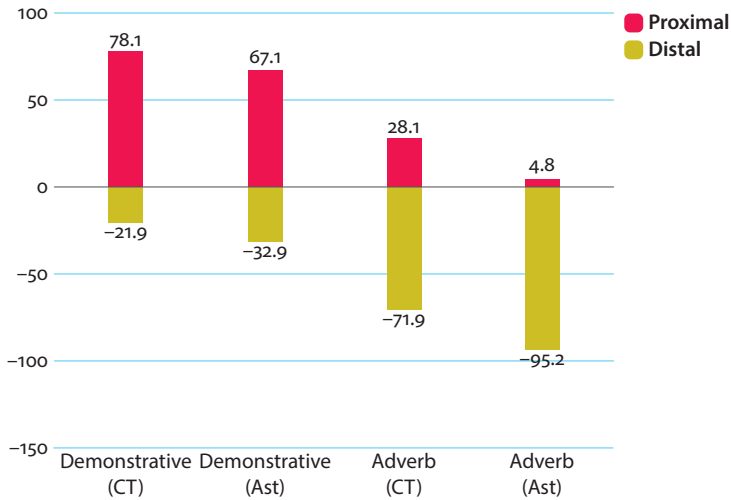
Figure 2. Proximal and distal demonstratives and spatial adverbs

in particular, take up such a high share (95.2%) because the speaker quite often points to locations in the hearer's domain when explaining how to use the astrolabe, as seen in Section 3.1.

The proximal–distal ratios of the temporal system are summarized in Figure 3:
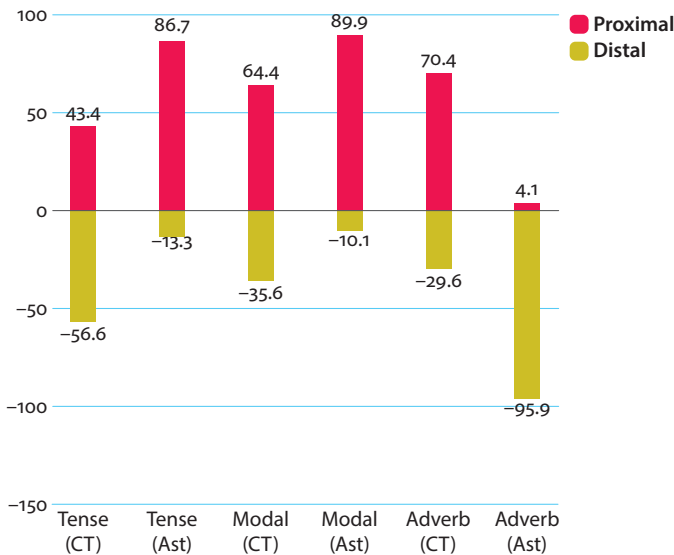
Figure 3. Proximal and distal tense forms, modals, and temporal adverbs

The ratio of the proximal vs. distal tenses strongly reflects the contrast between the simple present (proximal) and the simple past (distal). In CT, proximal and distal tense forms take up an almost equal share, with distal tense forms only slightly higher than proximal ones; in Ast, however, proximal tenses are employed much more frequently. The central modals are classified into proximal modals (*shall*, *will*, *can*, *may*, *must* [proximal]) and distal modals (*should*, *would*, *could*, *might*, *must* [distal]). Although proximal modals take up a higher share in both registers, this tendency is most marked in Ast. As already seen in 3.1, the speaker in Ast attends to scientific facts and relationships, while in CT, the narrators exploit a variety of senses of modals including distal senses (temporal, metaphorical, and hypothetical distance; see Nakayasu 2009, and also fn. 12 on the meanings and functions of modals). As regards temporal adverbs, the proportion of proximal and distal adverbs differs markedly in CT vs. Ast (see 3.1).

Figure 4 summarizes the preceding discussion by demonstrating which perspective, proximal or distal, the speaker is likely to take in CT and Ast.[21]

The proportions of proximal and distal elements above can be considered to reflect the difference in discourse structure of the texts. As seen in the introduction, CT is fiction, which has a complex discourse structure with four levels (Pakkla-Wekström 2005). It basically tells a past story, which is likely to promote the use of distal elements such as third-person pronouns and past tense forms. However, the speakers, i.e. Chaucer (Level 2) and the pilgrims (Level 3), take advantage of a wide variety of strategies:[22] for example, they use direct speech to describe the dialogue among the participants of the story (Level 4), they employ proximal elements to bring the hearers' attention to the story, and they frequently switch the

---

21. One might recall Biber's (1988) and Biber & Conrad's (2009) multidimensional (MD) analysis and note its similarities to the present paper's methodology. The MD analysis takes a quantitative approach in comparing many different registers. This approach analyzes co-occurrence patterns of various linguistic features and classifies these features into positive and negative ones (not evaluative but mathematical), which exhibit a complementary distribution. Here I carry out a quantitative analysis, classifying spatio-temporal elements as proximal or distal ones (assigned positive and negative percentage shares in the figures for the sake of convenience). Apart from the significant difference that the target of the present research is spatio-temporal systems, the critical difference is that the proximal and distal elements do not necessarily show strictly complementary distribution: proximal elements, or distal elements, may combine in a particular discourse fragment with other proximal elements only, or with other distal elements only, but they may also mix (see Section 4 for the qualitative analysis).

22. Biber & Conrad remarked (2009: 138–139) that "fiction is one of the most complex varieties to analyze from a register perspective, because the author must create a fictional world and can choose to describe that world from many different possible perspectives," and that "the variation concerns style, rather than register."
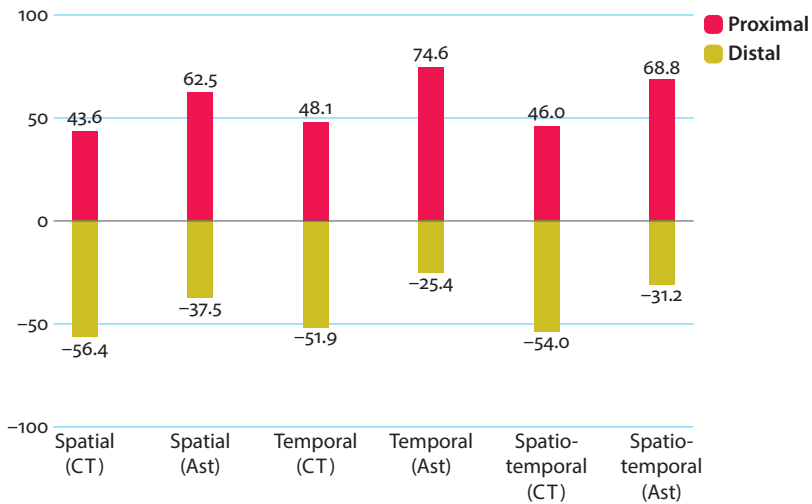
Figure 4. Summary of proximal and distal elements

perspective from distal to proximal (Section 4 returns to this). These strategies can explain the relatively high ratio of proximal elements in CT, although the ratio of distal elements, largely due to the past time frame, still remains slightly higher. On the other hand, in Ast, which is categorized as a scientific handbook, the discourse structure is not so complicated as in CT, and the speaker, Chaucer, addresses his son in Level 2. He describes scientific facts and relations often with the aid of proximal elements such as the present tense. In addition, the situation in which the speaker addresses the hearer promotes the use of proximal elements. For these reasons, the speaker is more likely to take the proximal perspective in both the spatial and temporal domains.

4.   Spatio-temporal systems in discourse

Having provided a quantitative description of the elements of space and time, this section conducts a qualitative analysis of spatio-temporal systems in Chaucer's language, further exploring the discourse structure of the two texts. The discussion first embarks on a micro-level analysis of typical combinations of spatio-temporal elements and then turns to elements of discourse structuring. An additional macro-level analysis demonstrates how proximal elements are combined to take a proximal perspective and distal elements to take a distal one, and how proximal and distal perspectives alternate with each other.

4.1   Typical combinations

As a first step, I examine some typical combinations of spatio-temporal elements and other related grammatical elements found in the corpus. First, the proximal demonstrative *this* and the text deictic element *forseide* "aforesaid" are combined to refer to an entity already mentioned in the text:[23]

(2)   14. *Than is there <u>a large pyn</u> in manere of an extre* ["axle tree"], *that goth thorugh the hole that halt the tables of the clymates and the riet* ["rete"][24] *in the wombe of the moder* ["mother"];[25] *thorugh <u>which pyn</u> ther goth a litel wegge* ["wedge"], *which that is clepid the hors, that streynith* ["holds"] *all these parties to-hepe* ["together"]. <u>*Thys forseide grete pyn*</u> *in manere of an extre is ymagyned to be the Pool Artik* ["the north celestial pole"] *in thyn Astralabie.* (Ast 1.14.1–9)

Here *thys forseide grete pyn* traces back – through *which pyn* – to *a large pyn*. The demonstrative *this* functions as an anaphoric element that draws the hearer's attention to the speaker's domain, thereby relating the referent of *thys forseide grete pyn* to *a large pyn*. This particular combination can only be found in Ast; 16 of 31 occurrences of *forseide* cooccur with the proximal demonstrative *this*/*these*. A similar anaphoric use of *this*, however, can also be observed in CT. Without combining with *forseide*, it appears with a character's name or their occupational title, as in *this Palamoun* (A.KN 1070); it thus refers to the person already mentioned in the previous context and directs the hearer's attention to that person again.[26]

   Imperatives may also select for specific spatio-temporal elements. In (3), after the successive use of imperatives *set* and *ley*, *than* "then" and the modal *shal* follow:

(3)   <u>*Set*</u> *the nadir*[27] *of thy sonne upon 18 degrees of height among thyn almykanteras* ["almucanters"][28] *on the west syde; and <u>ley</u> thy label on the degre of thy sonne, and <u>than</u> <u>shal</u> the point of thy label shewen the spryng of the day.* (Ast 2.6.1–5)

In the pattern 'imperative, and (*then*) proximal modal', the situation in the clause containing the proximal modal is realized immediately after the fulfillment of the

---

23. See Fries (1994) for the analysis of *foresaid* in Early Modern English.

24. "A planispheric projection of the sky" (Eisner 1976a: 130).

25. "The large disk that is the astrolabe proper" (Eisner 1975: 18).

26. Tanabe (2014) considers *this* as a pragmatic marker ('discourse marker' in this paper).

27. "A point on the celestial sphere diametrically opposite to another" (Benson 1987: 671).

28. "Circles of equal altitude parallel to the horizon" (North 1988: 53).

preceding condition/trigger; this pattern is attested to 19 times in the corpus.[29] Another example is the imperative occurring with the distal adverb *there*, as in (4):

(4) *<u>Set</u> the heved* ["head"] *of what signe the lyst to knowe his ascendyng in the right cercle upon the lyne meridionall* ["meridian (line)"],[30] *and <u>wayte</u> where thyn almury*[31] *touchith the bordure, and <u>set there</u> a prikke* ["dot"]; *<u>turne</u> than thy riet westward til that the ende of the forseide signe sitte upon the meridional lyne and eftsonys* ["soon after"] *<u>wayte</u> where thin almury touchith the bordure, and <u>set</u> <u>there</u> another pricke.* (Ast 2.27.1–9)

As already noted, the speaker is talking to his son who holds the astrolabe in his hands. When explaining how to use the astrolabe, the speaker refers to locations in the hearer's domain as *there*.

4.2 Discourse-structuring elements

This subsection examines the relationship between spatio-temporal systems and the elements which elaborate the structure of discourse. One such discourse-structuring element that was already discussed in 4.1 is anaphoric *this*, whose function derives from a proximal demonstrative, i.e. a spatial element. Other deictic elements that show a discourse-structuring function are time adverbs. Proximal *now*, for instance, which is used frequently in the corpus, is a temporal adverb that may function as a discourse marker. In (5), the Wife employs *now* to signal that she will change the topic to her fourth husband:

(5) *Swiche manere wordes hadde we on honde.*
   <u>Now</u> wol I speken of my fourthe housbonde.
   My fourthe housbonde was a revelour –
   This is to seyn, he hadde a paramour ["mistress"] *–*
   And I was yong and ful of ragerye ["playfulness"],
   Stibourn and strong, and joly as a pye ["magpie"]. (D.WB 451–456)

---

29. Nakayasu (2009: 220–224) notes that this kind of phenomenon is extensively observed in her Shakespeare corpus. The condition/trigger can be the *if*-clause, the imperative as in (3), the participial construction, the adverbial, and so on. See also example (6) and Nakayasu (2011).

30. "An imaginary line that passes from the poles through the zenith" (*MED* 2007).

31. "The pointer on the rete of an astrolabe at 0° Capricorn" (Benson 1987: 669).

Taavitsainen & Hiltunen (2012) argue that *now* introduces meta-comments as a discourse structuring device. It should be noted that together with *now*, the wife employs first-person pronouns and the proximal modal *wol*. *Now* cooccurring with first-person pronouns is also attested by Taavitsainen & Hiltunen (2012), and with the proximal modal *wol* by Nakayasu (2011: 89). The use of the discourse marker *now* in combination with proximal elements is a typical case where a situation is tightly connected to the speaker's domain.

The next temporal adverb that, like *now*, has a discourse structuring function is *thanne* "then". In its temporal meaning, it refers to a time in the past or the future; its constant temporal value is that of 'after a previous event' irrespective of whether the event is in the past or the future. It is often exploited by Chaucer, showing the order of events and a cause and effect relationship, as in (6):

(6) *Set the begynnyng of the degre that ascendith upon the ende of the 8 houre inequal* ["8th unequal hour"];[32] *than wol the begynnyng of the 2 hous* ["house"][33] *sitte upon the lyne of mydnight. Remeve* ["move"] *than the degre that ascendith, and set him on the ende of the 10 houre inequal, and than wol the begynnyng of the 3 hous sitte up on the mydnight lyne. …*   (Ast 2.36.1–7)

Like *now*, the adverb *than* is often accompanied by the imperative and a proximal modal (*wol* in (6)). The imperative in combination with *than* occurs 54 times in the corpus. Moreover, 19 out of 27 instances of the pattern 'imperative, *and than*' are used with a proximal modal, as in (3).

Among the elements which structure discourse, metadiscoursal phrases are particularly noteworthy. Metadiscourse is defined by Boggel (2009: 2) as "comprising all those elements in a text which do not add new material to the text proposition, but which are used to refer to already existing text-propositional elements." Since it is directly connected to the speaker's domain, it makes his presence overt in the text and exhibits how he interacts with the intended hearer. In (7), *this is to seyn* "this is to say" functions as metadiscourse:

(7) *Heigh labour and ful greet apparaillynge* ["preparation"]
     Was at the service and the fyr-makynge,
     That with his grene top the hevene raughte ["reached"];
     And twenty fadme of brede ["fathoms broad"] *the armes straughte* ["projected"] – *This is to seyn, the bowes weren so brode.*   (A.KN 2913–2917)

---

32. "The unequal hours are the time of daylight or night on any given day divided by 12" (Eisner 1975: 25).

33. "One of twelve astrological divisions of the heavens" (Benson 1987: 670).

Here all the tense forms are past, except for the present tense in the metadiscoursal phrase. While describing the story using the past tense, the speaker exploits the metadiscourse to structure the text, which is directly connected to the speaker's domain, and thus promotes the use of the present tense. A variety of metadiscoursal phrases can be found in the corpus, including *as I first seide* "as I first said" (Ast 1.6.9), *as I yow tolde* "as I told you" (A.KN 2097), *ther is namoore to seye* "there is no more to say" (A.KN 2366), *that is to seye* "that is to say" (D.WB 127), *as shal be shewid* "as shall be shown" (Ast 1.21.94), *I mene* "I mean" (A.KN 2216), and so on. In these phrases, the speaker may employ the past tense if he wants to refer to what he already said and the future form if he chooses to indicate what he plans to mention later.

4.3   Combining proximal or distal elements

Here we take a further step into a macro-level analysis and examine how the speaker controls spatio-temporal elements by combining proximal or distal elements and by alternating them in discourse. It is interesting to examine in this respect how consistently proximal elements and distal elements, are combined to achieve an overall proximal or distal perspective, respectively, or if, and to what extent, proximal and distal elements may co-occur but still evoke only a single perspective.

A first example of how proximal elements from different domains may be combined to yield a particular effect is (8), where the speaker Theseus talks to the knights Palamon and Arcite after intervening in their quarrel. He consistently combines proximal elements to evoke a proximal perspective, along with second-person pronouns to refer to these two knights:

(8)   *I yow foryeve al hoolly this trespaas,*
      At requeste of the queene, that kneleth heere,
      And eek ["also"] *of Emelye, my suster deere.*
      And ye shul both anon unto me swere
      That nevere mo ye shal my contree dere ["harm"],
      Ne make werre upon me nyght ne day,
      But been my freendes in all that ye may,
      I yow foryeve this trespas every deel.   (A.KN 1818–1825)

In addition to the first-person pronouns to refer to himself, the speaker employs proximal tense forms, proximal demonstratives, and proximal modals. Further, he employs a proximal spatial adverb *here* to draw his family into the situation and a temporal adverb *anon* to require the hearers to be temporally proximal to the speaker's domain, i.e. to take immediate action. It is likely that the intended

effect of this combination of spatial and temporal proximal elements is to boost the appeal made by the speaker to his audience. Moreover, he talks to the knights using second-person pronouns, which contributes to the speaker's proximity to the hearers' domain. This overall effect of immediacy involving both the speaker and the hearers provides the situation, so to speak, with an expanded proximal perspective (see 3.2).

The combination of interjections with other proximal elements offers another notable example of how an expanded proximal perspective may be achieved. Since interjections are strictly connected to the speaker's domain, they are strictly proximal and likely to be combined with other elements which are also related to the speaker's domain (and possibly the hearer's). Observe the following context in (9), where Arcite utters the interjection *What!* when he talks to his cousin Palamon:

> (9)   *What! Verray fool, thynk wel that love is free,*
>   *And I wol love hire maugree* ["in spite of"] *al thy myght!*   (A.KN 1606–1607)

In addition to the interjection, the speaker employs the address *verray fool* "true fool", an imperative, the present tense form, the first-person pronoun, a proximal modal, the second-person pronoun, and an exclamation. In a similar way to (8), these elements are either proximal to the speaker's domain or in the second-person's domain, and therefore they reinforce the overall proximal perspective attributed to the situation. The combination of these various elements supports the present paper's view that seemingly unrelated elements, including modals, may in fact constitute integrated spatio-temporal systems.[34]

The combination of distal elements can be observed particularly in narrations. The context in (10) describes what happens after Palamon has finished his prayer:

> (10)   *Whan the orison* ["player"] *was doon of Palamon,*
>   *His sacrifice he dide, and that anon,*
>   *Ful pitously, with alle circumstaunces* ["ceremony"],
>   *Al telle I noght as now his observaunces;*
>   *But atte laste the statue of Venus shook,*
>   *And made a signe, whereby that he took*
>   *That his preyere accepted was that day.*   (A.KN 2261–2267)

The narrator depicts the events as they happen, successively employing third-person pronouns and past tense forms, and also adverbials which motivate the use of the past tense (*whan the orison was doon of Palamon* "when the prayer of Palamon was done", *that anon* "that quickly", and *that day*) and specify the order of events

---

34. See also Example (13), where distal tense forms combine with a distal modal.

(*atte laste* "at last"). The two phrases containing a distal demonstrative *that* evoke a situation to be distanced. It should not be overlooked that Chaucer also employs proximal elements – i.e. first-person pronouns, the present tense form, and the adverb *now* – in *Al telle I noght as now* "Although I tell not now". Another interesting combination of distal elements with other elements which promote a distal perspective can be observed in the following fragment (11):

(11)   *The ascendent sothly* ["truly"], *to take it at the largest, is* <u>*thilke*</u> *degre that ascendith at eny of these forseide tymes upon the est orisounte* ["the observer's horizon"].[35] *And therfore, yf that eny planete ascende at* <u>*thatt same*</u> *tyme in* <u>*thilke*</u> *forseide degre, than hath he no latitude fro the ecliptik lyne* ["ecliptic line"],[36] *but he is than in the degre of the ecliptik which that is the degre of his longitude.* <u>*Men sayn*</u> *that* <u>*thilke*</u> *planete is in horoscopo. … For,* <u>*after the statutes of astrologiens*</u>, …
(Ast 2.4.7–19)

Here the coordination seems rather loose, since the pronouns are distal, whereas the tense forms and the demonstrative *these* are proximal.[37] The speaker is referring to third parties, i.e. *men* and *astrologiens* "astrologers"; this encourages the use of third-person pronouns and thus moves the perspective away from the speaker's domain.[38] Also motivating the distal perspective are such expressions as *thatt* and *same* – though the distalness of these expressions is not in fact related to that of the third parties. In addition, the employment of the demonstrative *thilke*, which is neutral as regards proximal or distal perspective, does not bring the speaker and the hearer back to their (proximal) domain but maintains the distal perspective.

4.4   The alternation between proximal and distal perspectives

In addition to the combination of proximal elements, or distal elements, attested in the corpus, a macro-level analysis of discourse also reveals that the speaker can alternate between proximal and distal perspectives. The discourse fragment in (12) offers an obvious clue to the change in perspective:

---

35. "The plane parallel to it through the center of the earth, extended to the celestial sphere" (Benson 1987: 663).

36. "The apparent path of the Sun through the fixed stars is a particularly important great circle on the celestial sphere" (North 1988: 24).

37. See Section 4.1 for the function of *forseide*.

38. The present tense still ties the third parties to the speaker's domain because they are either contemporaries or people from the past whose statements are still true at the time of speaking.

(12)   *Greet <u>was</u> the feeste in Atthenes <u>that day</u>,*
       And eek the lusty ["pleasing"] *seson of <u>that May</u>*
       <u>Made</u> every wight ["person"] *to been in swich plesaunce* ["delight"]
       That <u>al that Monday justen</u> they and <u>daunce</u>,
       And <u>spenden</u> it in Venus heigh servyse.   (A.KN 2483–2487)

The narrator starts depicting the day of a battle in a detached tone, by employing the past tense forms together with a temporal adverb phrase (*that day*) and a noun phrase with a distal demonstrative (*that May*). Once the distal spatio-temporal perspective has been set up, he employs the adverb phrase *al that Monday* to recast that perspective and switches the tense to the present to highlight what happens there although the time of the events is the past. This is an example of the historical present (tense), whose major function is to synchronize the story with the speaker's domain (and possibly the hearer's) (Benson 1961; Nakayasu 2013). Instances of the historical present are quite frequent in CT, where the narrator deliberately switches the temporal perspective from distal to proximal. In telling stories, he sometimes chooses to disregard the actual time reference and switches perspective to produce a dramatic effect.

The alternation between proximal and distal perspectives in discourse can be observed in either the spatial or the temporal domain, as well as in an integrated spatio-temporal domain. First, the alternation may occur only in the temporal domain. The use of the historical present as in (12) is an example where the alternation is used to make the reader connect to the past more. Alternatively, the tense form can also shift precisely to encode the actual time reference. The example in (13) shows a combination of *ensample* and the past tense:

(13)   *<u>Ensample as thus</u>: The yeer of oure Lord 1391, the 12 day of March at midday, I <u>wolde</u> knowe the degre of the sonne. I <u>soughte</u> in the bakhalf* ["reverse side"] *of myn Astrelabie and <u>fond</u> the cercle of the daies, the whiche I knowe by the names of the monthes writen under the same cercle. Tho <u>leyde</u> I my reule* ["rule"][39] *over this foreseide day, and <u>fond</u> the point of my reule in the bordure upon the firste degre of Aries, a litel within the degre. And thus knowe* ["solve"] *I this conclusioun* ["proposition"].   (Ast 2.1.6–16)

In this context, the speaker employs the past tense forms and the distal modal *wolde*[40] to describe the *ensample*. When he talks about what he obtains in the

---

39. "A straight-edge that pivots on the back of the astrolabe" (Eisner 1976b: 222).

40. The distal modal *wolde* here retains its original lexical meaning and signifies the speaker's past intention. Therefore, the distance from the speaker's domain is temporal, not metaphorical or hypothetical.

present, he switches to present tense: *And thus knowe I this conclusioun* "And thus I solve this proposition". The same strategy can be observed in (14), where the speaker shifts the tense from past to present in *thus have I 2 degres bitwixe bothe prikkes* "thus I have two degrees between the two points":

(14)   *Also the degre peraventure* ["by chance"] *of Jupiter, or of another planete,* _was_ *in the first degre of Piscis in longitude, and his latitude* _was_ *2 degrees meridional; tho* _tok_ *I the point of A and* _sette_ *it in the first degre of Piscis on the ecliptik; and than* _sette_ *I the point of F dounward in the same signe by cause that the latitude* _was_ *south 2 degres, that is to seyn, fro the heved of Piscis; and thus* _have_ *I 2 degres bitwixe bothe prikkes.*   (Ast 2.40.51–60)

Second, the alternation may be relevant only to the spatial domain. Consider the simple but clear contrast between the demonstratives *this* and *that*:

(15)   *Thy zodiak* ["zodiac"][41] *of thy Astrelabie is …. Amiddes* _this_ *celestial zodiak is ymagined a lyne which that is clepid the ecliptik lyne, under which lyne is evermo the wey of the sonne. Thus ben there 6 degres of the zodiak on* _that_ *oo syde of the lyne and 6 degrees on* _that_ *othir.* _This_ *zodiak is dividid in 12 principale divisiouns that departen the 12 signes, and, for the streitnesse* ["small size"] *of thin Astrolabie, than is every small divisioun in a signe departed by two degrees and two, I mene degrees contenyng 60 mynutes. And* _this_ *forseide hevenysshe zodiak is …*
                                             (Ast 1.21.32–50)

The speaker consistently employs the proximal demonstrative *this* to refer to the zodiac, but when he starts to describe the ecliptic line, he switches to the distal demonstrative *that*. Note that the two instances of *that* are followed by *oo* "one" and *othir* "other", respectively. The pattern '*that one* X + *that other* X' can be considered a complex distal determiner because *other* is itself already distal. He then returns to the zodiac once again with *this* so that his son can follow his explanation. Importantly, the speaker does not move his temporal perspective to and from proximal, because he continues to explain the astrolabe to his son face-to-face.

Finally, the proximal and distal perspectives may alternate in the integrated spatio-temporal domain. In (16), repeated from (1), Theseus explains to Palamon and Emelye about the god who created the universe, contrasting the *Firste Moevere* and the world he created:

---

41. "The imaginary celestial circular band or belt bisected by the ecliptic that contains all the positions occupied by the known planets and that is divided into twelve equal portions named after their regnant signs" (*MED* 2007).

(16)   *"The Firste Moevere of the cause above,*
       Whan <u>he</u> first <u>made</u> the faire cheyne of love,
       Greet <u>was</u> th'effect, and heigh <u>was his</u> entente. …
       <u>That</u> same Prince and <u>that</u> Moevere," <u>quod he</u>,
       "<u>Hath stablissed</u> in <u>this</u> wrecched world adoun
       Certeyne dayes and duracioun
       To al that <u>is</u> engendred in <u>this</u> place,
       Over the whiche day <u>they may</u> nat pace,
       Al <u>mowe they</u> yet <u>tho</u> dayes wel <u>abregge</u> …"   (A.KN 2987–2999)

Theseus adopts a distal perspective when he talks about the first mover, employing distal elements of both the spatial and temporal domains such as the distal demonstrative *that*, the third-person pronoun to refer to the first mover, past tense forms, and adverbials which promote the use of the past tense such as the *whan* "when" clause and *first*. Using the present perfect, he shifts his perspective down to the world the god made, *this wrecched world adoun* "this wretched world below", and starts using proximal elements integrating both spatial and temporal domains: the proximal demonstrative *this*, present tense forms, and the proximal modal *may* – an exception is the third-person pronoun *they*, which refers to all the people living in this wretched world. This vivid contrast between god and the world he created offers an excellent example of how the speaker manages the elements and the perspectives which constitute spatio-temporal systems in discourse.

5.   Conclusion

This paper has reported a systematic analysis of synchronic spatio-temporal systems in Chaucer's language against the background of research in historical pragmatics and discourse analysis. Based on a corpus consisting of a work of fiction, *The Canterbury tales*, and a scientific handbook, *A treatise on the astrolabe*, I provide a quantitative study of the elements/expressions of space and time employed by the speaker. The discourse structure of both registers is differently reflected in the selection of spatio-temporal elements. The speaker employed proximal elements more frequently in the scientific handbook where he addresses his young son; in contrast, the fictional work shows a complex discourse structure in which the speaker exploits a variety of strategies, correlating with a slightly higher frequency of distal elements.

A qualitative analysis of the discourse shows how spatio-temporal elements are interrelated with each other to take either a proximal or distal perspective and how these perspectives change in discourse. The micro-level analysis reveals some typical combinations of spatio-temporal elements and related elements such as deictic

expressions at the textual level and the imperative, and then the analysis highlights the discourse structuring elements, such as discourse markers and metadiscoursal phrases. Finally, the macro-level analysis of how the discourse is structured reveals how proximal elements are combined to convey a proximal perspective and, similarly, how distal elements are combined to convey a distal perspective. In addition, I have shown how the dynamic alternation between proximal and distal perspectives achieves a number of pragmatic effects, such as creating a dramatic effect in storytelling and highlighting the contrast between proximal and distal perspectives, e.g. *this world* and *that world* as in (16).

In conclusion, by focusing on the proximal–distal contrast across the various manifestations of such spatio-temporal elements as pronouns, demonstratives, adverbs, tense, and modals, I have provided insights into the homogeneity of what are often seen as unrelated phenomena.

The results of this study also raise new research questions. Further interesting results may be derived from incorporating into future analysis, for example, the relationship between the typical deictic use of a spatio-temporal element and its anaphoric use in the extended spatio-temporal domain in discourse, and the principles underlying extensions of meaning from the spatial to the temporal domain. In addition, comparisons with the systems in other registers such as religious texts will make the analysis more effective in highlighting the distinctive features of each system; this, then, may shed light on the synchronic spatio-temporal systems as a whole that develop in the Middle English period.

Acknowledgements

References

Benson, Larry Dean. 1961. Chaucer's historical present: Its meaning and uses. *English Studies* 42. 65–77. doi:10.1080/00138386108597094

Benson, Larry Dean (ed.). 1987. *The Riverside Chaucer*, 3rd edn. Boston: Houghton Mifflin.

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511621024

Biber, Douglas. 2004. Modal use across registers and time. In Anne Curzan & Kimberly Emmonds (eds.), *Studies in the history of the English language,* vol. 2: *Unfolding conversations*, 189–216. Berlin: Mouton de Gruyter. doi:10.1515/9783110897661.189

Biber, Douglas & Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511814358

Biber, Doublas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson Education.

Boggel, Sandra. 2009. *Metadiscourse in Middle English and Early Modern English religious texts: A corpus-based study*. Frankfurt a.M.: Peter Lang.

Brisard, Frank (ed.). 2002. *Grounding: The epistemic footing and deixis and reference*. Berlin: Mouton de Gruyter. doi:10.1515/9783110899801

Bühler, Karl. 1934. *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Jena: G. Fischer.

Coates, Jennifer. 1983. *The semantics of modal auxiliaries*. London: Croom Helm.

Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: John Benjamins. doi:10.1075/tsl.42

Duchan, Judith F., Gail A. Bruder & Lynne E. Hewett (eds.). 1995. *Deixis in narrative: A cognitive science perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Eisner, Sigmund. 1975. Building Chaucer's astrolabe, Part I. *Journal of British Astronomical Association* 86. 18–29.

Eisner, Sigmund. 1976a. Building Chaucer's astrolabe, Part II. *Journal of British Astronomical Association* 86. 125–132.

Eisner, Sigmund. 1976b. Building Chaucer's astrolabe, Part III. *Journal of British Astronomical Association* 86. 219–227.

Fillmore, Charles J. 1997 [1975]. *Lectures on deixis*. Stanford: CSLI Publications. [Originally published as *Santa Cruz lectures on deixis*. Indiana: Indiana University Linguistics Club.]

Fries, Udo. 1994. Text deixis in Early Modern English. In Dieter Kastovsky (ed.), *Studies in Early Modern English*, 111–128. Berlin: Mouton de Gruyter. doi:10.1515/9783110879599.111

Fritz, Thomas A. 2003. "Look here, what I am saying!" Speaker deixis and implicature as the basis of modality and future tense. In Friedrich Lenz (ed.), *Deictic conceptualisation of space, time and person*, 135–151. Amsterdam: John Benjamins. doi:10.1075/pbns.112.08fri

Gunther, Robert Theodore. 1977. *Chaucer on the astrolabe with original illustrations*. Revised from the 1931 Oxford Edition by Robert Theodore Gunther. Emeryville, CA: Norman Greene.

Halliday, Michael A. K. & Ruqaiya Hasan. 1976. *Cohesion in English*. Harlow: Longman.

Halliday, Michael A. K. & Christian M. I. M. Matthiessen. 2014. *Halliday's introduction to functional grammar*, 4th edn. Milton Park: Routledge.

Huang, Yan. 2014. *Pragmatics*, 2nd edn. Oxford: Oxford University Press.

Jucker, Andreas H. 2006. "Thou art so loothly and so oold also": The use of *ye* and *thou* in Chaucer's *Canterbury tales*. *Anglistik* 17(2). 57–72.

Jucker, Andreas H. & Irma Taavitsainen. 2013. *English historical pragmatics*. Edinburgh: Edinburgh University Press.

Kytö, Merja (ed.). 1996. *Manual to the diachronic part of the Helsinki corpus of English texts: Coding conventions and lists of source texts*, 3rd edn. Helsinki: Department of English, University of Helsinki.

Langacker, Ronald W. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780195331967.001.0001

Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Martin, James Robert. 1992. *English text: System and structure*. Amsterdam: John Benjamins. doi:10.1075/z.59

Martin, James Robert. 2015. Cohesion and texture. In Deborah Tannen, Heidi E. Hamilton & Deborah Schiffrin (eds.), *Handbook of discourse analysis*, vol. 1, 2nd edn., 61–81. Chichester: Wiley Blackwell.

MED. 2007. *Middle English dictionary*. http://quod.lib.umich.edu/m/med. (6 September, 2015.)

Nagucka, Ruta. 2000. The spatial and temporal meanings of before in Middle English. In Irma Taavitsainen, Terttu Nevalainen, Päivi Pahta & Matti Rissanen (eds.), *Placing Middle English in context*, 329–337. Berlin: Mouton de Gruyter. doi:10.1515/9783110869514.329

Nakayasu, Minako. 2009. *The pragmatics of modals in Shakespeare*. Frankfurt a.M.: Peter Lang.

Nakayasu, Minako. 2011. Towards a pragmatic analysis of modals *shall* and *will* in Chaucer's language. *Studia Anglica Posnaniensia* 46(4). 73–96. doi:10.2478/v10121-010-0012-7.

Nakayasu, Minako. 2013. Chaucer's historical present: A discourse-pragmatic perspective. In Liliana Sikorska & Marcin Krygier (eds.), *Evur happie and glorious, ffor I have at will grete riches*, 41–60. Frankfurt a.M.: Peter Lang.

Nakayasu, Minako. 2014. Weep now namoore; I wol thy lust fulfille: Towards a sociopragmatic analysis of the spatio-temporal systems in Chaucer. Paper read at the First Poznań Historical Sociopragmatics Symposium. Adam Mickiewicz University, Poznań, Poland.

Nakayasu, Minako. 2015. Spatio-temporal systems in *A treatise on the astrolabe*. In Juan Camilo Conde-Silvestre & Javier Calle-Martín (eds.), *Approaches to Middle English: Variation, contact and change*, 243–259. Frankfurt a.M.: Peter Lang.

North, John David. 1988. *Chaucer's universe*. Oxford: Clarendon Press.

Oakeshott-Taylor, John. 1984. Factuality, tense, intonation and perspective: Some thoughts on the semantics of 'think'. *Lingua* 62. 289–317. doi:10.1016/0024-3841(84)90056-1

Pakkala-Wekström, Mari. 2005. *The dialogue of love, marriage and* maistrie *in Chaucer's* Canterbury tales. Helsinki: Société Néophilologique.

Palmer, Frank Robert. 2001. *Mood and modality*, 2nd edn. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139167178

Taavitsainen, Irma. 1999. Personality and styles of affect in the *Canterbury tales*. In Geoffrey Lester (ed.), *Chaucer in perspective: Middle English essays in honour of Norman Blake*, 218–234. Sheffield: Sheffield Academic Press.

Taavitsainen, Irma & Turo Hiltunen. 2012. *Now* as a text deictic feature in Late Medieval and Early Modern English medical writing. In Ulrich Busse & Axel Hübler (eds.), *Investigations into the meta-communicative lexicon of English: A contribution to historical pragmatics*, 179–205. Amsterdam: John Benjamins. doi:10.1075/pbns.220.14taa

Taavitsainen, Irma & Andreas H. Jucker. 2010. Trends and developments in historical pragmatics. In Andreas H. Jucker & Irma Taavitsainen (eds.), *Historical pragmatics*, 3–30. Berlin: De Gruyter Mouton.

Taavitsainen, Irma & Andreas H. Jucker. 2015. Twenty years of historical pragmatics. *Journal of Historical Pragmatics* 16. 1–24. doi:10.1075/jhp.16.1.01taa

Tanabe, Harumi. 2014. *Kantaberiimonogatari* no katarite no shakaikaikyuu kara mita goyoo-hyooshiki toshite no *this* ni tsuite [The pragmatic marker *this* viewed from the narrator's social rank in *the Canterbury tales*]. In The Society for Chaucer Studies & Koichi Kano (eds.), *Choosaa to chuusei o nagamete: Choosaakenkyuukai 20shuunenkinenronbunshuu* [Through the eyes of Chaucer: Essays in celebration of the 20th anniversary of the Society for Chaucer Studies], 185–200. Tokyo: Asao Press.

Traugott, Elizabeth Closs. 1972. *A history of English syntax: A transformational approach to the history of English sentence structure*. New York: Holt, Rinehart and Winston.

Traugott, Elizabeth Closs. 1974. Explorations in linguistic elaboration: Language change, language acquisition and the genesis of spatio-temporal terms. In John Mathieson Anderson & Charles Jones (eds.), *Historical linguistics*, vol. 1: *Syntax, morphology, internal and comparative reconstruction: Proceedings of the First International Conference on Historical Linguistics, Edinburgh, 2–7 September 1973*, 263–314. Amsterdam: North-Holland Publishing Company.

Traugott, Elizabeth Closs. 1978. On the expression of spatio-temporal relations in language. In Joseph H. Greenberg (ed.), *Universals of human language*, vol. 3: *Word structure*, 369–400. Stanford: Stanford University Press.

Wilkins, David P. 1995. Expanding the traditional category of deictic elements: Interjections as deictics. In Judith F. Duchan, Gail A. Bruder & Lynne E. Hewitt (eds.), *Deixis in narrative: A cognitive science perspective*, 359–386. Hillsdale, NJ: Lawrence Erlbaum.

"A riddle to myself I am"

Argument shifting in English congregational song between 1500 and 1900

Kirsten Gather

Heidelberg University

This article presents and refines the main results of a larger study on left-shifted arguments in congregational song (see Gather 2014) and discusses its implications for genre studies in general. Argument shifting[1] is defined as the deviation from any unmarked order of obligatory arguments, as in the title of the famous carol *We three kings of Orient are*. Very often, requirements of meter and rhyme are evident causes of shifting. In this article, I provide the results obtained for congregational song, which lead to two major patterns of argument shifting with distinct syntactic and poetic criteria. The comparison to secular poetry shows the uniqueness of the phenomenon in congregational song and the conservatism of the genre.

Keywords: diachronic syntax, argument shift, word order, corpus linguistics, genre, congregational song, metrical psalm, hymn

1. Introduction

Congregational songs were an integral and characteristic part of English church services for several centuries. Although they altered in contents from sung versifications of biblical material to freely authored religious poetry of a personal and sometimes intimate nature, they largely maintained their rather strict poetic form. It is these rigid formal requirements of meter and rhyme in particular which led to deviations from standard contemporary word order, as, for instance, in the song titles *He who would **valiant** be* and *All people that **on earth** do dwell*.

---

1. In this study, the terms *shifting* and *moving* are applied without a particular syntactic theory in mind since argument shifting in congregational songs is not syntactically motivated.

The present study investigates different kinds of argument shifting in English congregational songs from the beginnings of the genre in the 16th century until its gradual decline towards the end of the 19th century.[2] A breakdown into syntactic functions of shifted arguments shows that shifted objects are the most frequent and most characteristic subtype.

The focus of the analysis is on the interaction between poetic requirements and syntactic features such as the internal phrase structure of the shifted argument and its position in the clause. I show that certain poetic and syntactic factors are likely to co-occur. This leads to two major patterns of argument shifting, which are manifested most clearly in shifted objects and which depend on their distribution within the predominant subgenre of congregational song. To assess the results, I then look at object shifting in secular poetry and compare the findings for congregational song and secular poetry.

The main hypothesis is that argument shifting, and especially object shifting, is a relevant and characteristic feature of congregational song, which makes the genre archaic in comparison to other verse genres. I argue in particular that the distribution of the two patterns of argument shifting contributes to the genre's conservatism. This also allows conclusions with regard to the conservative nature of congregational song in comparison to other genres in the religious domain.

The study has the following parts. First, since the various forms of song have never been looked at as one genre, I give a definition of *congregational song* (Section 2). Next, I introduce argument shifting and the different kinds of core arguments that are shifted (Section 3). Section 4 briefly comments on the corpus and methodology of my study. Section 5 presents the main results of the analysis, discussing the two major patterns of argument shifting in congregational song. A comparison with the data obtained from secular verse shows that this phenomenon is significant. Finally, I also discuss the results against the background of other religious genres and genre studies in general. I argue that argument shifting is an additional characteristic that makes congregational song conservative not only compared to secular verse but also in comparison to religious prose genres.

---

2. This study is based on a larger study of argument shifting in congregational songs (Gather 2014). Besides including the main results presented here, the larger study also analyzes the phenomenon in more detail, looking at authors' personal styles and at the connection between text and music. The present study refines the main results of Gather (2014), and discusses them in comparison with other genres in the religious domain and in the context of genre studies in general.

2.   English congregational song as a genre

The term *congregational song* is used as an umbrella term for several subtypes of liturgical song. It can best be described by its characteristic features, which distinguish the genre from other related genres.

English congregational song is, to start with, metrical and rhymed English poetry in stanzaic form. In contrast to secular poetry, it usually addresses the Christian Deity, a Saint, or fellow Christians. Moreover, in contrast to most other poetry, it is designed to be sung and it is (meant to be) performed by a congregation of people, usually as part of a liturgical service.

The meter that is most frequently applied in congregational songs is the so-called *common meter*, a specific sequence of iambic tetrameters and trimeters. The majority of rhyme schemes are either ABCB or ABAB.

As regards the existence of source texts, congregational songs can be divided into songs with biblical origin and those containing freely authored texts. The most frequent subtype of songs with biblical origin is the *metrical psalm*, followed by the *canticle* and other versified biblical material. The category of congregational songs with non-biblical origins consists exclusively of *hymns*.[3]

Congregational songs with a biblical origin, and metrical psalmody in particular, constitute the predominant subtype of the genre in the 16th and 17th centuries. From the second half of the 17th century onwards, hymns, i.e. congregational songs with non-biblical texts, by and by take over and become the predominant subgenre. The main reason for the transition from biblical to non-biblical texts is the growing opposition to metrical psalmody and its archaic diction (see, for instance, Barton 1656; Patrick 1679).

Figure 1 illustrates the different subtypes of congregational song (in bold print).
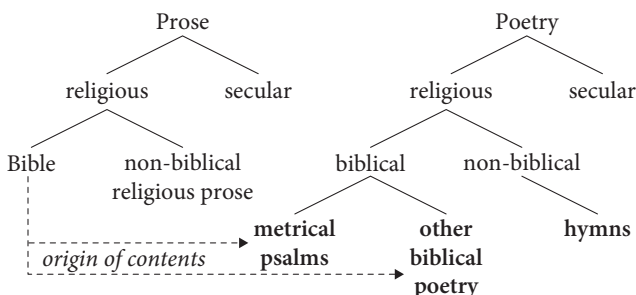
Figure 1. Schematic overview of the subgenres of congregational song

---

3.   What is meant here are *traditional hymns*, not the hymns that came up in the second half of the 20th century (see Section 4).

Passages (1) and (2) illustrate the two main subtypes of congregational song, metrical psalms and hymns.

(1)   *Then faithfully calle thou on me*
      in dayes of thy distresse:
      I will thee free and thou my name
      shalt praise with thankfullnesse.   Henry Dod (1620), *Psalm 50*

In (1), which is taken from a metrical psalm, God speaks to the people of Israel. The stanza shows the common meter of metrical psalms, a sequence of iambic tetrameters and trimeters, and the typical rhyme scheme found in common meter, which is ABCB.

Hymns are less restricted in their topics than songs with biblical origin. Their topics are usually more personal and intimate than those of metrical psalmody. Often, sinners beseech the Holy Trinity to redeem them from their suffering, as in the following stanza from Wesley.

(2)   *Eternal Sun of Righteousness,*
      Display thy beams divine!
      And cause the glories of thy face
      Upon my heart to shine.   Charles Wesley (1780), *Hymn CCXLIII*

In terms of poetic requirements, hymns offer more variety in rhyme scheme and meter. Within a single hymn, however, poetic requirements are usually adhered to. The rhyme scheme in (2) is traditional (common meter, rhyme scheme ABAB).

3.   Argument shifting

After this brief overview and illustration of the genre *congregational song*, it is essential to define and categorize what I call *argument shifting*.

The general assumption is that, at least from the 16th century onwards, unmarked argument order, that is, SV(X), is already rather fixed (for instance, Kohonen 1978: 123–125; Rissanen 1999: 264),[4] so that any kind of deviation from SV(X) that is not grammatically required happens for a reason. In the case of poetry, the reasons are usually metrical requirements or constraints due to the rhyme scheme (see Görlach 1991: 35). As is shown in Sections 5 and 6, different kinds of

---

4.   Although linguists also discuss arguments in marked positions in Early Modern English texts, corpus-based studies are usually restricted to prose genres where arguments in marked positions form the exception rather than the rule (for instance, Speyer 2010: 27, who shows that percentages of topicalized objects are below 5% in Early Modern English).

deviations from unmarked word order in congregational song can help determine the degree of conservatism of the genre in comparison to other verse genres.

Before looking at argument shifting in detail, note that I focus on obligatory clause elements since optional arguments can be placed basically anywhere in the clause (see Quirk et al. 1985: 490). Obligatory clause elements comprise objects, subject and object complements, and certain adverbials (Quirk et al.: 55–57). Subjects are not part of the investigation because there are hardly any cases of subject shifting.

Argument shifting, then, is understood as the deviation of objects, subject/object complements, or obligatory adverbials from their unmarked positions in the clause, unless syntactically required, for instance, in certain types of question or relative clause. Usually, the argument is shifted to the left, which means that it occurs either left of the subject and the full verb, e.g. XS(Aux)V, or between the subject and the full verb, e.g. SX(Aux)V.

Argument shifting in the history of English has been discussed under different terms depending on the position of the shifted argument. If the argument is shifted to initial position, this is usually known as *topicalization* or *fronting* (e.g. Fischer et al. 2004; Speyer 2010). If it is shifted to medial position between subject and verb, linguists commonly refer to *OV order* instead of *VO order* (e.g. van der Wurff & Foster 1997; Moerenhout & van der Wurff 2005).

The change from verb-final to non-verb-final order in Middle English times should, however, be seen in a wider context than that of the OV/VO alternation. The fixation of non-verb-final word order is the common denominator of the markedness of argument shifting. Whether the argument is moved to initial or to medial position is secondary because both cases are considered 'improper word order' by contemporaries (e.g. Wither 1624; Barton 1656; Patrick 1679) and both kinds of shifting are usually triggered by the same causes, i.e. poetic requirements. Also, discussions of verb final and non-verb-final word order should not be restricted to the object since the change can be attested for subject/object complements and obligatory adverbials, too. Kohonen (1978: 123–125), for instance, shows that SXV order in general began to break down in Early Middle English and that it was the complement which moved beyond the verb first, not the object.

A first illustration of argument shifting in congregational songs is the following passage from a metrical psalm by Francis Seager.

(3)   *For thou dost **them** defende and saue*
      From threates of myghty poure:
      ***From venym tounges** thou dost **them** hyde*
      Wythin thy pleasaunt houre.   Francis Seager (1553), *Psalm 31*

In this example, *thou* (referring to God) is the one who defends and saves *them*, i.e. those who fear God, and the one who hides them from venomous tongues. The passage has two shifted pronominal objects (*them*), which are placed between auxiliary and main verb, and an adverbial in initial position (*From venym tounges*), which usually follows the main verb. In the case of pronoun shifting, metrical requirements often play a role, rather than rhyme. *For thou dost them defend and saue* is an iambic tetrameter, whereas the regular argument order, *for thou dost defend and saue them*, has no regular scansion. By contrast, the initial position of the fronted adverbial in line three is due to neither meter nor rhyme. The line could as well have read *Thou dost them hyde from venym tounges*. It is likely that Seager intended to emphasize the opposition of the wicked, who are characterized by their 'venom tongues'. Apart from fronting for emphasis, Seager applies a figure of speech here. Since the end of the first clause and the beginning of the second clause are adverbials of the same kind, the two clauses form a chiasmus:

$$S\,Aux\,O\,V\,A$$
$$\diagdown\!\!\diagup$$
$$\diagup\!\!\diagdown$$
$$A\,S\,Aux\,O\,V$$

Although the use of rhetorical devices is usually secondary to maintaining meter and rhyme, it nevertheless helps explain those cases in which neither meter nor rhyme trigger argument shifting.

A second example is the following fragment from Frances Ridley Havergal's hymn *To thee*.

(4) ***My heart to Thee*** *I bring,*
The heart I cannot read;
A faithless, wandering thing,
An evil heart indeed.
I bring it, Saviour, now to Thee,
*That **fixed and faithful** it may be.*   Frances Ridley Havergal (1884), *To thee*

In Havergal's hymn as well, there are several elements in marked positions: the object *my heart* is fronted, the obligatory adverbial *to thee* occurs before subject and verb, and, in the last line, the subject complement *fixed and faithful* is in initial position, instead of *that it may be fixed and faithful*.[5]

---

5. *To thee* is categorized as a space adjunct. For a detailed explanation of why space adjuncts are classified as adverbials and not as prepositional objects, see Gather (2014: 28–29). A general discussion of this multiple analysis can be found in Quirk et al. (1985: 1150–1167). The term

Here, the first two instances of argument shifting are due to the maintenance of the rhyme scheme since only *bring* rhymes with *thing*. The third instance happens for reasons of meter and rhyme because unmarked argument order would not have iambic scansion, nor would it rhyme with the line before.

As shown in Section 5, (3) and (4) are typical instances of two different patterns of argument shifting which are characterized by distinct poetic and syntactic features. The first pattern consists of shifted pronominal objects in medial position between subject and (full) verb. These shifts often occur due to metrical constraints and are almost exclusively found in congregational songs with biblical origin. There are almost no instances in hymns. In Seager's stanza (Example (3): *them*), lines one and three contain instances of this first pattern.

The second argument shifting pattern is characterized by longer and heavier arguments in clause-initial position. The purpose of these shifts is usually to maintain the rhyme scheme. A typical instance of the second pattern is found in the first line of Havergal's stanza, with an object in fronted position (Example (4): *My heart*).

To summarize, I investigate the left-shifting of obligatory clause elements to marked positions in the clause. These obligatory arguments are objects (e.g. John Cennick's *O my Dear GOD, **my Strength** renew*), complements (e.g. George Wither's *And **blest** is he, that died for thy sake*), and obligatory adverbials (e.g. Mary Sidney Herbert's *Your service cheerfully **on him** imploy*).

I look at both syntactic criteria and poetic requirements in the analysis. Syntactic criteria are the phrase structure of the shifted argument, i.e. the word classes involved, the position of the shifted argument with regard to the other obligatory clause elements, and the influence of auxiliaries. The phrase structure of the shifted argument is relevant in connection with its weight and prominence in the clause. The main distinction in the position of the argument is whether it appears initially (e.g. XSV) or medially (e.g. SXV) in the clause. The presence or absence of auxiliaries is relevant because they, and especially dummy *do*, are often used to provide an extra syllable in the line. This increases the number of positions an argument can occur in (e.g. SAuxXV, SXAuxV).

Poetic requirements are constraints of meter and rhyme. In a few cases, figures of speech are also relevant. But when both poetic requirements and rhetorical devices are involved, the first usually overrule the latter. Requirements of meter and rhyme would never be given up in order to employ a figure of speech, and there are many instances where only poetic requirements can explain argument shifting.

Before turning to results, the next section provides basic information on the corpus and methodology.

'complement' is used in a rather narrow sense in Quirk et al. (1985: 54–55). It comprises what other grammarians call 'predicative complements' or 'predicative adjuncts' (see Bussmann 2006: 377).

4.   Corpus and methodology

In this section, I explain how the corpus of congregational songs was compiled, the criteria of analysis, and the terminology employed. I furthermore briefly describe the cross-section of secular poetry to which the results of congregational song are compared.

It makes sense to start investigating congregational song with the beginning Reformation in England, i.e. after 1500, because the Reformation marked the start of congregational participation as such (see Temperley 1979: 37). The year 1900 indicates the endpoint of the corpus because afterwards there is not much new hymn material apart from the so-called 'hymn explosion' of the 1960s and 1970s. These new hymns, however, are songs of a completely different character than the 'traditional' church song. They have simple and repetitive texts, do not draw on any hymn tradition, and are influenced by popular music instead (see Temperley 2001: 33). Already in the 19th century it becomes more and more difficult to find enough new material since many hymnbooks are compilations of older hymns, often without naming authors or dates of origin.

I selected 15 authors per century with as much variety in denomination, profession, and gender as possible, and their songbooks range from the most popular and widespread books of their time to song collections that were meant for only one small parish church. From every songbook, I chose songs randomly until I reached a word count of roughly 2,000 words.[6] I avoided the doubling of psalms in the choice of songs, i.e. the doubled selection of one and the same psalm number, in order not to skew the results.

Thus, the corpus comprises 60 samples of roughly 2,000 words each, with 15 samples per century:

Table 1. Corpus of congregational songs: Number of texts, number of words

	16th century	17th century	18th century	19th century	Total
Number of texts	15	15	15	15	60
Number of words	30,023	30,006	30,004	30,011	120,044

To be able to interpret the results, I compiled a smaller corpus for comparison. It contains a cross section of Early and Late Modern English secular poetry (ca. 40,000 words), again with 2,000 word samples. The focus of the selection was on the diversity of texts. While the original study (Gather 2014) focused on Early Modern texts for comparison, the present study compares congregational song with secular

---

6.   The appendix contains a list of the texts and the detailed results of the analysis. The same applies to the samples of secular poetry.

poetry across the entire Modern English period and therefore makes use of a more comprehensive dataset.

Table 2. Corpus of Early and Late Modern English secular poetry: Number of texts, number of words

	16th century	17th century	18th century	19th century	Total
Number of texts	5	5	5	5	20
Number of words	9,977	10,004	10,056	10,173	40,210

The electronic corpus texts are manual transcriptions. The sources are *Early English books online (EEBO)*, *Eighteenth century collections online (ECCO)*, *Nineteenth century collections online (NCCO)*, and paper editions. Analyses were carried out manually. Full bibliographical information on individual texts is provided in the references section.

It is not sufficient to work with normalized frequencies in order to make developments in the data visible. The number of shifted arguments per 10,000 words does not always lead to conclusive results if it is not seen against the background of all obligatory arguments that may be moved. For instance, if samples A and B both contain 40 shifted objects but sample A has 120 objects that could be shifted and sample B only has 80 (all are normalized frequencies), then the ratio of shifted to 'shiftable' arguments is much more relevant than normalized frequencies, which for samples A and B would be equal.[7]

Therefore, I counted the numbers of objects, complements, and obligatory adverbials that could be shifted and all shifts that actually occurred and calculated their ratios. For every instance of argument shifting, I coded the order of the core arguments (shifted argument in initial or medial position, in particular), the grammatical categories of the shifted clause element, the presence or absence of auxiliaries, and whether the shift is due to meter, rhyme, or both or was not due to any perceptible reason. Table 3 below lists the syntactic and poetic parameters analyzed.

Table 3. Syntactic and poetic parameters under investigation

Parameters	Values
position of shifted argument	initial or medial
internal phrase structure	PRN, D (A) N, etc.[8]
presence of auxiliaries	yes or no
poetic requirements	rhyme, meter, both or neither

---

7. A similar approach can be found in Speyer (2010: 27), who calculates the relative share of topicalized direct objects as the percentage of all sentences with direct objects.

8. PRN = pronoun, D = determiner, A = adjective, N = noun.

Correlation analyses then helped reveal the connection between the three subtypes of argument shifting and find parallel or inverted developments of the syntactic and poetic features. Chi-square tests were applied.[9]

5.   Argument shifting in English congregational song

In the following, I first present the main results of the analysis of argument shifting in congregational song. In a second step, I show the influence of meter and rhyme on the phenomenon and discuss the kinds of arguments that have shifted. I focus on shifted objects, since they are by far the most frequently moved clause elements. I show that there are two major patterns of object shifting and that the distribution of these patterns reflects the development of the genre.

5.1   Main results

Table 4 provides the number of arguments that may be shifted, the number of shifts that actually occur, and the proportion of actually shifted arguments to shiftable arguments.

Table 4. Argument shifting in English congregational song
(absolute frequencies and percentages)[10]

	16th century	17th century	18th century	19th century	Total
Number of words	30,023	30,006	30,004	30,011	120,044
Number of obligatory arguments that may be shifted	3,597	3,639	3,521	3,066	13,823
Number of shifted arguments	1,274	1,471	924	779	4,448
Ratio	35.4%	40.4%	26.2%	25.4%	32.2%

It turns out that argument shifting is indeed rather prominent in congregational song: at least every fourth, and on average every third, argument is moved. The high percentage of shifted arguments is surprising because any kind of deviating word order is already considered marked in the 16th century (see Section 3). As is shown in Section 6, the percentage of shifted arguments is much lower in other verse genres, such that this is a characteristic feature of congregational song.

---

9.   For the correlation analysis, I use Pearson's correlation coefficient (Wessa 2015) with normalized frequencies. On Chi-square tests, see Preacher (2001).

10.   The data in Table 4 are statistically significant ($\chi^2 = 128.779$; $df = 3$; $p < 0.001$).

Interestingly, the share of shifted arguments decreases suddenly from the 17th to the 18th century, from about 40% to roughly 26%. As is shown below, this correlates with the change of the dominant subgenre from metrical psalms to hymns.

5.2 Meter and rhyme

It is clear from the examples that meter and rhyme play an important role with respect to argument shifting in congregational song. To determine whether meter and/or rhyme are relevant factors triggering argument shifting in a clause, we can transform the clause that contains the shifted argument into its unmarked equivalent. If constraints of meter or rhyme are violated, we can assume that argument shifting happened for these reasons, amongst others. This leads to Table 5 below:

Table 5. Meter and rhyme as reasons for argument shifting
(absolute frequencies and percentages)[11]

Reason for argument shifting	16th century	17th century	18th century	19th century	Total
Meter and rhyme	540	690	467	390	2,087
	42.4%	46.9%	50.5%	50.1%	46.9%
Rhyme only	412	467	377	306	1,562
	32.3%	31.8%	40.8%	39.3%	35.1%
Meter only	195	206	44	47	492
	15.3%	14.0%	4.8%	6.0%	11.1%
Neither meter nor rhyme	127	108	36	36	307
	10.0%	7.3%	3.9%	4.6%	6.9%
Total	1,274	1,471	924	779	4,448
	100%	100%	100%	100%	100%

Table 5 shows that, indeed, meter and rhyme can be assumed as triggers of argument shifting in at least 90% of the cases. Moreover, meter and rhyme in combination trigger the majority of all shifts, and meter-only shifts decrease in favor of rhyme-only and meter-and-rhyme shifts. As in Table 4, there is an obvious break between the 17th and the 18th centuries,[12] with meter-only cases decreasing to roughly 5% and cases of shifting due to (meter and) rhyme rising to over 90%.

---

11. Aggregating the data of the first two rows (rhyme and rhyme-and-meter cases) and rows three and four (meter cases and shifts for neither meter nor rhyme) and aggregating the data of the first two centuries and the last two centuries shows that the break around 1700 is significant ($\chi^2 = 131.873$; $df = 1$; $p < 0.001$).

12. It should be noted, however, that periodization by century obscures when exactly the actual change took place. Nevertheless, the temporal connection between this break and the transition from psalmody to hymnody around 1700 is very likely (see Section 5.4).

5.3   Subtypes of argument shifting

Let us now take a closer look at the arguments that are shifted. I distinguish between shifted objects, complements, and obligatory adverbials (see Quirk et al. 1985: 49–57).

Table 6. Argument shifting by syntactic function (absolute frequencies and normalized frequencies per 1,000 words)[13]

	16th century	17th century	18th century	19th century	Total
Shifted objects	820	879	638	476	2,813
	27.3	29.3	21.3	15.9	23.4
Shifted complements	124	204	53	83	464
	4.1	6.8	1.8	2.8	3.9
Shifted adverbials	330	388	233	220	1,171
	11.0	12.9	7.8	7.3	9.8
Total	1,274	1,471	924	779	4,448
	42.4	49.0	30.8	26.0	37.1

Table 6 and Figure 2 present the frequency distribution of all shifts per century, split up by syntactic function of the shifted argument. Apparently, more than half of all instances per century are objects. In fact, changes in the normalized frequencies
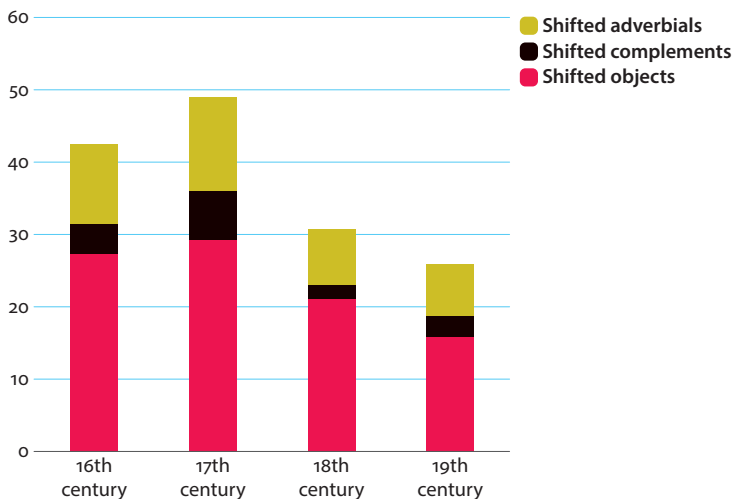
Figure 2. Shifted objects, complements, and obligatory adverbials (normalized frequencies per 1,000 words)

---

13. The data in Table 6 are statistically significant (16th to 17th centuries: $\chi^2 = 12.171$; $p = 0.0023$; 17th to 18th centuries: $\chi^2 = 43.007$; $p < 0.001$; 18th to 19th centuries: $\chi^2 = 18.336$; $p < 0.001$).

of shifted objects correlate strongly with changes in the frequencies of all shifted arguments.[14]

As shown in Gather (2014: 183–190), all three subtypes of argument shifting display similar developments of syntactic and poetic features. Because objects are the most frequent of the three argument types and due to their strong correlation with the development of all shifted arguments, they provide a solid basis for examining the major patterns of argument shifting more closely.[15]

5.4   The two major patterns of argument shifting

Focusing on objects as the most frequently shifted argument, let us now look at the sudden decrease in object shifting at the turn of the 18th century. As mentioned in 5.2, this decrease goes along with a decrease in meter-only shifting in favor of shifting due to rhyme.

Table 7. Parallel developments in four object shifting parameters (absolute frequencies and percentages of all shifted objects per century)

	16th century	17th century	18th century	19th century
Internal phrase structure: PRN	250/820 30.5%	171/879 19.5%	48/638 7.5%	34/476 7.1%
Object in medial position (SOV, etc.)	447/820 54.5%	436/879 49.6%	266/638 41.7%	175/476 36.8%
Clause with shifted object contains auxiliary	444/820 54.1%	442/879 50.3%	213/638 33.4%	158/476 33.2%
Meter-only shifting	126/820 15.4%	117/879 13.3%	26/638 4.1%	24/476 5.0%

As seen in Table 7 and Figure 3, the sudden decrease in shifted objects in the 18th century (see also Table 6) correlates with a decline in the proportions of shifted pronouns (abbreviated as PRN in the table), objects shifted to medial position, auxiliary involvement in shifted elements, and meter-only shifting. These four features can be seen as constituting a pattern which is quite prominent in congregational songs of the 16th and 17th centuries but which is used much less frequently after 1700. This pattern, which is illustrated in (5), is not limited to metrical psalms and can be traced back to Middle English poetry, where it is quite frequent (see Gather 2014: 204). Chaucer, for instance, uses it a lot.

---

14. Pearson's correlation coefficient is 0.9770 (two-sided $p = 0.0230$).

15. For shifted complements, in particular, it would be necessary to enlarge the corpus considerably in order to achieve clearer results. The correlation coefficient of shifted complements and all shifted arguments is 0.8912 ($p = 0.1088$).
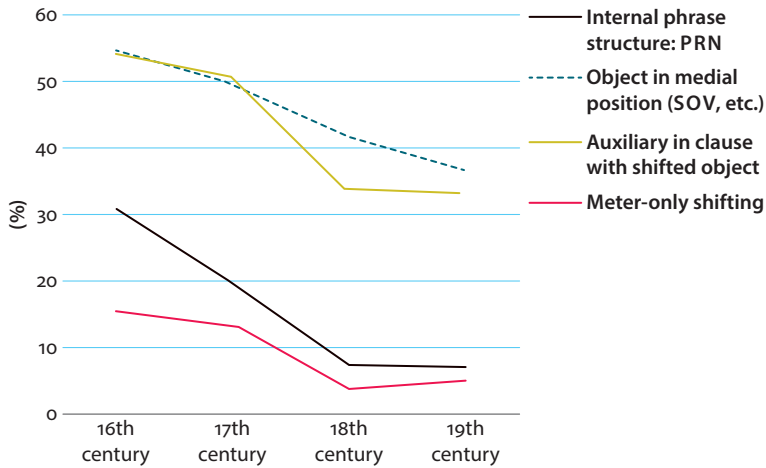
Figure 3. Parallel developments in four object shifting parameters (percentages of all shifted objects per century)

(5)   *Thou mad'st me feel affliction sore,*
      *and yet thou didst **me** save;*
      *Yea, thou didst help, and **me** restore,*
      and took'st me from the grave.

                    Thomas Sternhold and John Hopkins (1562), *Psalm 71*

The decrease in shifted objects in the 18th century also marks the transition to another pattern of object shifting. As can be seen from Table 8 and Figure 4, the percentage of shifted objects with a more complex internal phrase structure, usually [determiner (adjective) noun], rises, as does that of objects in initial position in the clause. As well, shifts due to rhyme increase, and the involvement of auxiliaries decreases. Unlike the first pattern of object shifting, this second pattern thus contains longer and heavier objects in clause-initial position that are shifted due to rhyme, and without the presence of auxiliaries in the clause.

Table 8. Parallel developments in object shifting (absolute frequencies and percentages of all shifted objects)

	16th century	17th century	18th century	19th century
Internal phrase structure: D (A) N ( )	422/820 51.5%	520/879 59.2%	497/638 77.9%	361/476 75.8%
Object in initial position	370/820 45.2%	432/879 49.1%	372/638 58.3%	300/476 63.0%
No auxiliary in clause with shifted object	376/820 45.9%	437/879 49.7%	425/638 66.6%	318/476 66.8%
Rhyme-only shifting	245/820 29.9%	298/879 33.9%	277/638 43.4%	208/476 43.7%
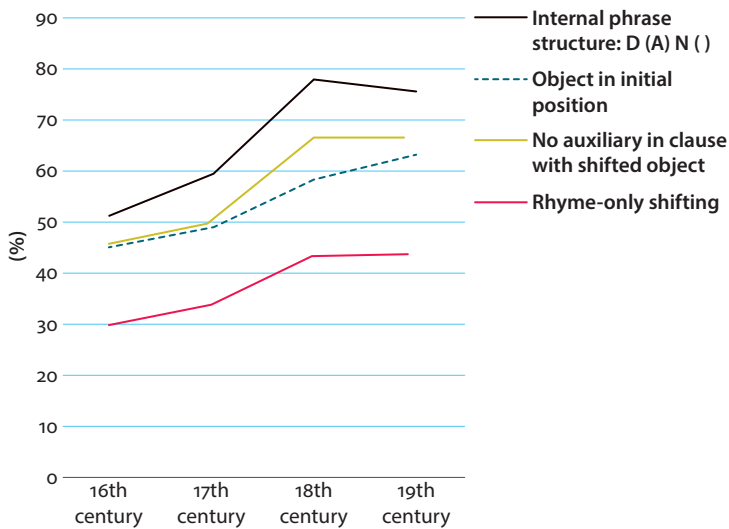
Figure 4. Parallel developments in object shifting (percentages of all shifted objects)

Table 8 and Figure 4 show the similar developments of the various features that form this second pattern.

The stanza in (6), taken from a hymn by James Montgomery, contains two examples of this second pattern in succession.

(6) *The wounded and the weak*,
He comforts, heals, and binds;
The lost He came from heaven to seek,
And saves them when He finds.   James Montgomery (1826), *Hymn 478*

Apart from poetic requirements, this kind of shifting is often used for reasons of emphasis rather than to hide the shifted argument in the middle of the clause. When long and heavy arguments are shifted to initial position in the clause (as, for instance, in Hammond's song title *A riddle to myself I am* in the title of this paper), this appears to be done deliberately, the more likely so because there is a free choice of words; in contrast, the first pattern may have been applied out of the necessity to be as faithful to the biblical source as possible.

The following table summarizes the features of the two major patterns of argument shifting.

Table 9. The features of the two major patterns of object shifting

	First pattern	Second pattern
Internal phrase structure	PRN	D (A) N
Position of shifted argument	medial	initial
Presence of auxiliaries	yes	no
Poetic requirements	meter only	rhyme only

Given that the above sets of characteristic features form the two most prominent patterns of argument shifting, what are their relative shares? If only pure instances are counted, i.e. instances where all characteristics apply, what are the proportionate shares of first and second patterns?[16] Table 10 shows frequencies[17] and percentages.

Table 10. Object shifting: Instances of first (I) and second (II) patterns in congregational song (absolute frequencies, normalized frequencies (per 10,000w), and relative shares in percentages)[18]

Congregational song	16th century		17th century		18th century		19th century		Total	
	I	II	I	II	I	II	I	II	I	II
Absolute frequency	198	223	116	145	19	262	8	205	341	835
Normalized frequency	65.9	74.3	38.7	48.3	6.3	87.3	2.7	68.3	28.4	69.6
Proportion I/II (in %)	47.0	53.0	44.4	55.6	6.8	93.2	3.8	96.2	29.0	71.0

As can be seen in Figure 5, the two patterns are roughly in balance in the 16th and 17th centuries. After 1700, the second pattern is clearly dominant although there are still a number of instances of pattern I remaining. Taking the change of subgenres around 1700 into account (see Section 2), it becomes clear that the first pattern is closely linked to the subgenre of metrical psalms while it occurs only rarely in hymns. Pattern II, in contrast, has been present in the genre of metrical psalms from the beginning, but it became dominant with the advent of hymnody and the end of metrical psalmody.

However, despite the second pattern becoming more prominent with the end of Bible-based congregational song, it should be noted that pattern II does not sufficiently compensate the loss of pattern I, such that the overall frequency of shifted objects decreases in the 18th and 19th centuries (see Table 6).

It is a matter of speculation why the first pattern is retained so long in metrical psalms. Possibly, this retention is connected to the authors' attempted faithfulness to the biblical source text and to the relatively high frequency of pronouns in the Psalter. The latter results from the very limited number of involved participants (e.g. a praying person [*I*] addressing God [*thou*] or talking about the wicked [*they*]). Such pronouns are often shifted for reasons of meter and rhyme, since words such

---

16. This approach can only be a rough approximation because there are many cases in which one feature is missing. Nevertheless, some general statements on the distribution and development of the two patterns can be made.

17. I include normalized frequencies here in order to make the data for congregational song comparable to the results obtained for secular poetry in Section 6.

18. Aggregating the frequencies of the first two and the last two centuries, a Chi-square test for patterns I and II yields $\chi^2 = 229.089$; $df = 1$; $p < 0.001$.
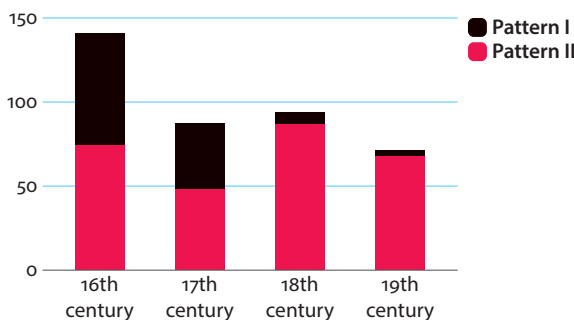
Figure 5. Object shifting: instances of first and second patterns in congregational song (normalized frequencies)

as *me*, *thee*, or *them* are usually avoided in the stressed position at line end. Instead, they are frequently inconspicuously placed between subject and verb, so that the verb can be used for end-rhyme.

In hymns, however, pronouns are often replaced by noun phrases because there is no need to be faithful to any source text. God, for instance, becomes *the helper of the helpless* (Wesley 1780), and the addressor is not *I* or *we* but *thy servant* (Cennick 1741/1742) or *our willing souls* (Doddridge 1755).

5.5   Summary

To sum up, argument shifting is a notable phenomenon in congregational song. On average, every third object, complement, or obligatory adverbial is moved to a marked position. Most instances of shifting are subject to the constraints of meter and/or rhyme.

An investigation of the type of argument that is shifted in congregational song reveals that object shifting is the most frequent subtype, and, within that subtype, two major shifting patterns of congregational song can be discerned. The older, archaic pattern, already found in Middle English poetry, consists of short shifted arguments, usually pronouns, in medial position in the clause and shifted due to metrical requirements. The second pattern, which already occurs at the beginning of English congregational song but becomes predominant only in hymnody, is characterized by longer, heavier arguments in initial position in the clause, which are shifted in order to maintain the rhyme scheme.

In sum, the main reason for the sudden decrease in argument shifting in congregational song after 1700 was the loss of the first shifting pattern with the end of bible-based congregational song and the insufficient compensation by the second shifting pattern.

6.   Comparison: Argument shifting in secular poetry

Given the results presented in Section 5, the question arises whether these particular manifestations of argument shifting – with the overall high ratios of shifted arguments, the sudden decrease after 1700, and the distribution of the two patterns – are unique to congregational song or not. Similar patterns might occur in other types of poetry, such as in secular verse. If so, a different explanation may be necessary to account for their loss in the 18th century. To test to what extent the findings in the previous section are unique to congregational song as a genre, I have also examined the frequency of shifted arguments and the distribution of the shifting patterns in a selection of secular poems from 1500 to 1900.

Table 11 and Figure 6 present the proportion of shifted objects to 'shiftable' objects, both in congregational song and in the cross section of secular verse.

Table 11. Object shifting in congregational song and in secular poetry (absolute frequencies and percentages of all shiftable objects)[19]

	16th century	17th century	18th century	19th century
Congregational song: shifted objects vs. shiftable objects	820/2,279	879/2,268	638/2,467	476/1,944
Percentage	36.0%	38.8%	25.9%	24.5%
Secular poetry: shifted objects vs. shiftable objects	266/1,015	158/924	109/679	62/534
Percentage	26.2%	17.1%	16.1%	11.6%

What these figures clearly show is that the percentages of shifted objects (relative to all shiftable objects) in secular poetry *are* considerable but that they are always well below the percentages in congregational song: 25% of shifted objects in secular poetry in the 16th century roughly equal the percentages in congregational song in the 18th and 19th centuries, after object shifting had decreased considerably.

---

19. Comparing the behavior of congregational song and secular poetry in the first two and the last two centuries by means of a Chi-square test yields $\chi^2 = 70.869$; $df = 1$; $p < 0.001$.
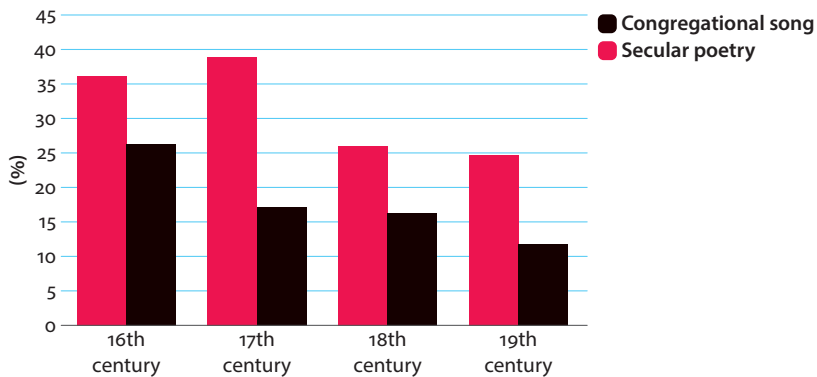
Figure 6. Object shifting in congregational song and in secular poetry (as a percentage of all shiftable objects)

Table 12 and Figures 7a (repeating Figure 5) and 7b, then, compare the relative shares of pattern I and pattern II.

Table 12. Instances of the first (I) and second (II) patterns in congregational song and secular poetry (absolute frequencies, normalized frequencies [per 10,000w] and relative shares in percentages)[20]

	16th century		17th century		18th century		19th century		Total	
Congregational song	I	II	I	II	I	II	I	II	I	II
Absolute frequency	198	223	116	145	19	262	8	205	341	835
Normalized frequency	65.9	74.3	38.7	48.3	6.3	87.3	2.7	68.3	28.4	69.6
Proportion I/II (in %)	47.0	53.0	44.4	55.6	6.8	93.2	3.8	96.2	29.0	71.0
Secular poetry										
Absolute frequency	43	60	2	24	0	32	0	16	45	132
Normalized frequency	43.1	60.1	2.0	24.0	0.0	31.8	0.0	15.7	11.2	32.8
Proportion I/II (in %)	41.7	58.3	7.7	92.3	0.0	100	0.0	100	26.5	73.5

---

20. Analogous to the approach for congregational song (see fn. 17), a Chi-square test for patterns I and II in secular poetry yields $\chi^2 = 19.451$; $df = 1$; $p < 0.001$. Comparing congregational song and secular poetry yields $\chi^2 = 3.831$; $df = 1$; $p = 0.0503$ for pattern I and $\chi^2 = 23.342$; $df = 1$; $p < 0.001$ for pattern II.
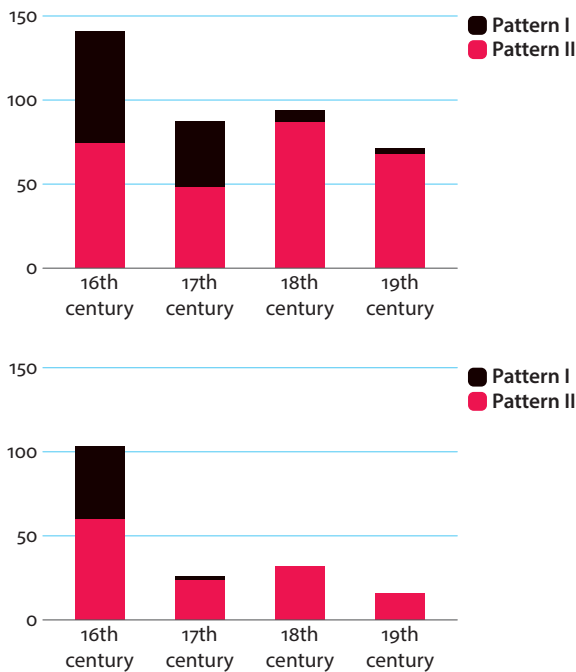
Figures 7a and 7b. Instances of the first and second patterns in congregational song (above) and secular poetry (below), normalized frequencies (per 10,000w)

Figure 7a shows the distribution of the two patterns in congregational song. As previously observed, the second pattern, i.e. longer arguments in clause-initial position and shifting triggered by rhyme, takes over only after 1700, and there are still a few instances of pattern I in the 18th and 19th centuries. In secular poetry (see Figure 7b), the second pattern already becomes predominant in the 17th century, i.e. a century earlier, and there are no remaining instances of the first pattern in the secular poetry of the 18th and 19th centuries. This indicates that congregational song lags behind in shedding pattern I and promoting pattern II as the only possible one, while secular verse is (as one might expect) more progressive.

    To sum up, object shifting in secular verse is much less frequent. As in congregational song, there is a transition from the first to the second pattern. The first pattern, however, dies out in secular poetry a century earlier than in congregational song, which seems to lag behind in this respect.

7.   Conclusion

What do the above findings say about congregational song as a genre, when, on the one hand, comparing them with findings from secular poetry and other religious genres, and, on the other hand, considering them in the context of genre studies in general?

As shown in 5.1, every third obligatory argument is shifted in congregational song on average. There is a sudden decrease of the percentage of moved arguments after 1700, mainly due to the almost complete loss of the first, archaic pattern along with the end of metrical psalmody and the insufficient compensation by the second pattern. The different distributions of the two patterns in metrical psalmody and hymnody suggest that the genre *congregational song* consists of two text types: biblical and non-biblical congregational song.[21] While the genre is situationally defined, the two text types differ in their linguistic characteristics.

Focusing on object shifting, the percentage of shifted objects is considerably lower in secular poetry than in congregational song, and the first shifting pattern vanishes a century earlier. This means that congregational song with its higher ratios and the longer retention of the first pattern is conservative, if not archaic, in comparison to secular poetry.

To what extent is this relevant for congregational song as a religious genre? Without going too much into detail, it can be said that congregational song contains many linguistic features which show its conservative character in comparison to other religious genres like sermons or religious treatises. Conservative features are, for instance, the use of *thou* instead of *you*, the third person singular verb ending *-eth* instead of *-s*, the use of *be* instead of *are*, and *the which* instead of *which* (see Kohnen 2011; Kohnen et al. 2011).

Let us consider the following example.

(7) *My sonne, yf thou wilt receaue my wordes,*
     and kepe my comaundementes by the,
     that thine eare maye herken vnto wysdome,
     applie thine herte then to vnderstandynge.

                              Matthew Bible (1537), Proverbs 2.1–2

---

21. On the distinction between *genre* and *text type*, see Biber (1993: 244–245).

(8)   *My son receiue ye these my wordes,*
      <u>*the which*</u> *shalbe right wise*
      And kepe thou my commaundementes,
      my son I **the** *aduise.*
      So that thine eares may euermore
      to wysedomes scholes *enclyne:*
      Applye thyne herte to vnderstand
      suche thynges as <u>*be*</u> *diuyne.*   Henry Howard (1550), Proverbs 2.1–2

This is a passage taken from the proverbs. Example (7) is the biblical source passage for Henry Howard's versification (Example 8). The two instances of argument shifting are in bold print. Where the biblical source has *thou* (*wilt*), *thee*, and *thine*, the versification contains the same archaic pronouns but additionally *the which* instead of *which* and *be* instead of modern *are*, both without equivalents in the Matthew Bible. This is one of many examples in which congregational song seems to accumulate the features that make religious genres archaic.

According to Kohnen (2011) and Kohnen et al. (2011), the genres in the religious domain form a continuum with very progressive genres, such as religious biographies, at one end and very conservative genres, such as prayers, at the other. Without going into detail, the diction of many congregational songs contains the same archaic morphological and syntactic features that are found in very conservative religious prose, such that congregational song ranks at the conservative end in this regard.

In the context of genre studies, argument shifting apparently constitutes an additional feature to assess whether a genre is conservative or progressive when comparing verse genres. Thus, congregational song is not only conservative in comparison to other religious genres, but it also inherits an archaizing component from verse genres, which it retains longer than secular poetry. This additional characteristic makes congregational song even more conservative than any other religious genre.

To conclude, this study shows that linguistic analyses of genre studies can indeed be informed by the study of formal characteristics only available to poetry, such as the constraints of meter and rhyme; in other words, verse genres should not be automatically excluded from linguistic analysis. In this study, peculiarities of word order which are commonly dismissed as poetic license were exemplarily shown to act as an indicator of the progressiveness or conservatism of a genre.

Acknowledgements

Primary sources[22]

Corpus of congregational songs

Ainsworth, Henry. 1612. *The Book of Psalmes: Englished both in Prose and Metre. With Annotations, opening the words and sentences, by conference with other scriptures*. Amsterdam. STC 2407.

Allen, James & Christopher Batty. 1757. *A Collection of Hymns for the Use of Those that Seek, and Those that Have Redemption in the Blood of Christ*, 1st edn. Kendal: Thomas Ashburner.

Austin, John (= William Birchley) . 1668. *Devotions in the Ancient Way of Offices*. Paris. Wing A4248A.

Barton, William. 1644. *The Book of Psalms in Metre. Close and Proper to the Hebrew: Smooth and Pleasant for the Metre: Plain and easie for the Tunes*. London. Wing B2401.

Baxter, Richard. 1692. *Mr Richard Baxter's Paraphrase on the Psalms of David in Metre, with other Hymns*. London. Wing B2580.

Byrom, John. 1773. *Miscellaneous Poems*, 2 vols. Manchester: Joseph Harrop.

Cennick, John. 1741/1742. *Sacred Hymns for the Children of God in the Days of Their Pilgrimage*. London: B. Milles.

Chandler, John. 1837. *The Hymns of the Primitive Church, now first Collected, Translated, and Arranged*. London: John W. Parker.

Conder, Josiah. 1824. *The Star in the East*. London: Thomas Davison.

Coverdale, Miles. 1535. *Goostly Psalmes and Spirituall Songes*. London. STC 5892.

Cowper, William & John Newton. 1779. *Olney Hymns*. London: W. Oliver.

Crowley, Robert. 1549. *The Psalter of Dauid newely translated into English metre in such sort that it may the more decently, and wyth more delyte of the mynde, be reade and songe of al men*. London. STC 2725.

Dod, Henry. 1620. *All the Psalms of David*. London. STC 2731.

Doddridge, Philipp. 1755. *Hymns Founded on Various Texts in the Holy Scriptures*. London: Joshua Eddowes & John Cotton.

Drayton, Michael. 1591. *The Harmonie of the Church*. London. STC 7199.

Ellerton, John. 1888. *Hymns Original and Translated*. London: Skeffington.

Elliott, Charlotte. 1859 [1839]. *Morning and Evening Hymns for a Week*. London: Lionel Booth.

Faber, Frederick William. 1849. *Jesus and Mary, or, Catholic Hymns for Singing and Reading*. London: James Burns.

Fawcett, John. 1782. *Hymns Adapted to the Circumstances of Public Worship and Private Devotion*. Leeds: Griffith Wright & Son.

Gibbons, Thomas. 1784 [1769]. *Hymns adapted to Divine Worship*. London: Printed for James Buckland & Charles Dilly.

Gill, Thomas Hornblower. 1868. *The Golden Chain of Praise*. London: William Hunt & Co.

Hall, John. 1565. *The courte of vertue*. London. STC 12632.

Hammond, William. 1745. *Psalms, Hymns, and Spiritual Songs*. London: William Strahan.

---

22. Information on the electronic resources used may be found in Section 4. In some references Short Title Catalogue (STC) and Wing numbers are given. STC and Wing are part of the English Short Title Catalogue (ESTC), which lists publications from the 15th to 18th centuries.

Hart, John. 1759. *Hymns Composed on Various Subjects*. London: John Everingham.

Havergal, Frances Ridley. 1884. *Poetical Works*. New York/Chicago: Fleming H. Revell.

Heber, Reginald. 1827. *Hymns, written and adapted to the Weekly Church Service of the Year*. London: John Murray.

Howard, Henry. 1550. *Certayne Chapters of the Prouerbes*. London. STC 2760.

Hunnis, William. 1583. *Seuen Sobs of a Sorrowfull Soule for Sinne. Comprehending those seuen Psalmes […] called Poenitentiall: framed into a forme of familiar praiers, and reduced into meeter by William Hunnis*. London. STC 13975.

Keach, Benjamin. 1691. *Spiritual Melody, Containing near Three Hundred Sacred Hymns*. London. Wing K93.

Keble, John. 1827. *The Christian Year*. London: Humphrey Milford, Oxford University Press, and Longmans, Green & Co.

King, Henry. 1651. *The Psalmes of David from the New Translation of the Bible turned into Metre*. London. Wing B2457.

Lok, Henry. 1597. *Ecclesiastes, otherwise called The Preacher. Containing Salomons Sermons or Commentaries […] vpon the 49. Psalme of Dauid his father*. London. STC 16696.

Lynch, Thomas Toke. 1855. *The Rivulet*. London: Longmans, Green, Reader, & Dyer.

Mason, John. 1691 [1683]. *Spiritual Songs or Songs of Praise to Almighty God upon several Occasions. Together with the Song of Songs*. London. Wing M92B.

Merrick, John. 1765. *The Psalms, Translated or Paraphrased in English Verse*. Reading: John Carnan & Co.

Montgomery, James. 1826. *The Christian Psalmist*. Glasgow: Printed for William Collins.

Neale, John Mason. 1843. *Hymns for the Sick*. London: Joseph Masters.

Parker, Matthew. 1567. *The whole Psalter Translated into English Metre*. London. STC 2729.

Patrick, John. 1679. *A Century of Select Psalms, and Portions of the Psalms of David, especially those of Praise. Turned into Metre, and Fitted to the Usual Tunes in Parish Churches*. London. Wing B2536.

Reeve, John. 1682. *Hymnes and Spiritual Songs extracted from Scripture*. London. Wing R672.

Samuel, William. 1569. *An Abridgement of all the Canonical Books of the Olde Testament, written in Sternholds meter*. London. STC 21690.

Seager, Francis. 1553. *Certayne Psalmes select out of the Psalter of Dauid*. London. STC 2728.

Sidney Herbert, Mary Countess of Pembroke, & Sir Philip Sidney. 1599. *The Psalmes of David translated into divers and sundry kinds of verse, more rare and excellent for the Method and Varietie than ever hath been done in English. Begun by the noble and learned Gent. Sir Philip Sidney, Knt. and finished by the right honourable The Countess of Pembroke, his sister*. First printed edition from original manuscript. London: Chiswick Press, 1823.

Smyth, Miles. 1668. *The Psalms of King David Paraphrased, And turned into English Verse, according to the common Metre, As they are usually sung in Parish Churches*. London. Wing B2494.

Southwell, Robert. 1595. *Moeoniae Or Certaine Excellent Poems and Spirituall Hymnes*. London. STC 22955.

Steele, Anne (= Theodosia) . 1780. *Poems on Subjects Chiefly Devotional*, 3 vols. Bristol: William Pine, James Buckland, & Joseph Johnson.

Sternhold, Thomas & John Hopkins (et al.) . 1562. *The Whole Booke of Psalmes ('Old Version')*. London. STC 2430.

Swain, Joseph. 1792. *Walworth Hymns, by J. Swain, Pastor of the Baptist Church Meeting There*. London: Printed for the author.

Tate, Nahum & Nicholas Brady. 1696. *New Version of the Psalms of David*. London. Wing B2606.

Thring, Godfrey. 1874. *Hymns and Sacred Lyrics*. London: Henry S. King & Co.

Toplady, Augustus Montague. 1759. *Poems on Sacred Subjects*. London: Printed for Edward & Charles Dilly.

Verstegan, Richard. 1601. *Odes In imitation of the seauen penitential psalmes, vvith sundry other poemes and ditties tending to deuotion and pietie*. Antwerp. STC 21359.

Waring, Anna Laetitia. 1854. *Hymns and Meditations*. London: W. & F.G. Cash.

Watts, Isaac. 1707. *Hymns and Spiritual Songs*. (Enlarged and corrected edition: 1709.) London: John Humfreys.

Wedderburn, John. 1578. *Ane co[m]pendious [buik] of godlie psalmes and spirit[uall sangis] collectit furthe of sindrie partis of the Scripture […]*. Edinburgh. STC 29667.

Wesley, Charles & John Wesley. 1780. *A Collection of Hymns, for the Use of the People called Methodists*. London: John Paramore.

Wither, George. 1623. *The hymnes and songs of the church diuided into two parts: the first part comprehends the canonicall hymnes, and such parcels of Holy Scripture, as may properly be sung, with some other ancient songs and creeds: the second part consists of spirituall songs, appropriated to the seuerall times and occasions, obseruable in the Church of England / translated and composed by G.W.* London. STC 25908.

Woodford, Samuel. 1697. *Select Hymns taken out of Mr Herbert's Temple, and turn'd into the common metre*. London. Wing H1515A.

Wordsworth, Christopher. 1865. *The Holy Year*. London: Rivingtons, Waterloo Place.

Wyatt, Sir Thomas. 1549. *Certayne Psalmes Chosen out of the Psalter of Dauid*. London. STC 2726.

Secular poetry

Blake, William. 1789, 1794. *Songs of Innocence and of Experience*. Copy C. Library of Congress, Electronic Edition. http://www.blakearchive.org/copy/songsie.c?descId=songsie.c.illbk.01 (12 December, 2015).

Browning, Robert. 1842. *The Pied Piper of Hamelin. A Child's Story*. London.

Byron, George Gordon Lord. 1800–1805. *Hours of Idleness and other Early Poems*. Reprint, Newcastle upon Tyne: Cambridge Scholars Publishing 2009.

Cowper, William. 1773–1784. Miscellaneous Poems. In Humphrey Sumner Milford (ed.). 1905. *The Complete Poetical Works of William Cowper*, 263–432. London: Henry Frowde.

Dryden, John. 1682. *MacFlecknoe. A Satyr upon the Trew-Blue-Protestant Poet T.S.* London: Printed for D. Green.

Hardy, Thomas. 1898. *Wessex Poems*. London: Harper.

Howard, Sir Henry & Thomas Wyatt (et al.) . 1557. *Songs and Sonnettes ('Tottel's Miscellany')*. Reprint from 1870, London: Edward Arber.

Johnson, Samuel. 1820. Poems. In Murphy, Arthur (ed.). *The Works of Samuel Johnson LL.D.*, vol. 1, 137–261. London: John Sharpe.

Jonson, Benjamin. 1640–1641. A Celebration of Charis in Ten Lyric Pieces. In *The workes of Benjamin Jonson*, vol. 2, 165–172. London: Richard Bishop. STC 14754.

Milton, John. 1667. *Paradise Lost*. London: Samuel Simmons.

More, Sir Thomas. 1516. A merry Jest. In Arthur Cayley, Jr. (ed.), *Memoirs of Sir Thomas More*, vol. 1, 34–48. Reprint from 1808, London: Cadell & Davis.

More, Sir Thomas. undated. Pageant Verses. In Arthur Cayley, Jr. (ed.), *Memoirs of Sir Thomas More*, vol. 1, 11–15. Reprint from 1808, London: Cadell & Davis.

Pope, Alexander. 1720–1730. *The Poetical Works of A. Pope*, vol. 1. Reprint from 1836, Philadelphia: Joseph J. Woodward.

Shakespeare, William. 1609. *Shakespeare's Sonnets*. London. STC 22353.

Shelley, Percy B. 1817. *The Revolt of Islam: A Poem in Twelve Cantos*. London: Printed for Charles & James Ollier by McMillan.

Skelton, John. 1545 [1521–1522]. *Colin Clout*. London. STC 22601.

Spenser, Edmund. 1579. *The Shepheardes Calender […]*. London. STC 23089.

Spenser, Edmund. 1590. *The Faerie Queene* […]. London. STC 23080.

Tennyson, Alfred Lord. 1830. The Early Poems. In John Churton Collins (ed.). 1900. *The Early Poems of Alfred, Lord Tennyson* […], 1–271. London: Methuen.

Waller, Afred Rayney. 1905–1906. *The English Writings of Abraham Cowley*, 2 vols. Cambridge: Cambridge University Press.

Wordsworth, William & Samuel T. Coleridge. 1798. *Lyrical Ballads*. Bristol.

References

Barton, William. 1656. *A View of many Errors and som gross Absurdities in the Old Translation of the Psalms in English Metre; as also in som other Translations lately published: Shewing how the Psalms ought to be translated, to be acceptable and edifying.* London. Wing B1007.

Biber, Douglas. 1993. Representativeness in corpus design. *Literary and Linguistic Computing* 8(4). 243–257. doi:10.1093/llc/8.4.243

Bussmann, Hadumod. 2006. *Routledge dictionary of language and linguistics*. Translated and ed. by Gregory P. Trauth & Kerstin Kazzazi. London: Routledge.

Early English Books Online (EEBO). http://eebo.chadwyck.com (1–23 December, 2015.)

Eighteenth Century Collections Online (ECCO). http://gdc.gale.com/products/eighteenth-century-collections-online (1–23 December, 2015.)

Fischer, Olga, Ans van Kemenade, Willem Koopman & Wim van der Wurff. 2004. *The syntax of Early English*. Cambridge: Cambridge University Press.

Gather, Kirsten. 2014. *Syntactic dislocation in English congregational song: A corpus-based study*. Frankfurt a.M.: Peter Lang. doi:10.3726/978-3-653-03805-7

Görlach, Manfred. 1991. *Introduction to Early Modern English*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139166010

Kohnen, Thomas. 2011. Religious language in 17th-century England: Progressive or archaic? In Joachim Frenk & Lena Steveker (eds.), *Anglistentag 2010: Proceedings*, 279–287. Trier: Wissenschaftlicher Verlag Trier.

Kohnen, Thomas, Tanja Rütten & Ingvilt Marcoe. 2011. Early Modern English religious prose: A conservative register? In Paul Rayson, Sebastian Hoffmann & Geoffrey Leech (eds.), *Methodological and historical dimensions of corpus linguistics* (Studies in Variation, Contacts and Change in English 6). Helsinki: University of Helsinki, VARIENG website. http://www.helsinki.fi/varieng/journal/volumes/06/kohnen_et_al/ (24 November, 2015.)

Kohonen, Viljo. 1978. *On the development of English word order in religious prose around 1000 and 1200 AD*. Åbo: Stiftelsens för Åbo Akademi.

Matthew, Thomas (= John Rogers) (ed.). 1537. *The Byble which is all the holy Scripture: in whych are contayned the Olde and Newe Testament truly and purely translated into Englysh by Thomas Matthew. Set forth with the Kinges most gracyous lycence* (= Matthew Bible). Antwerp: STC 2066.

Moerenhout, Mike & Wim van der Wurff. 2005. Object-verb order in early sixteenth-century English prose: An exploratory study. *English Language and Linguistics* 9(1). 83–114. doi:10.1017/S1360674305001553

Nineteenth century collections online (NCCO). http://gale.cengage.co.uk/product-highlights/history/nineteenth-century-collections-online.aspx (1–23 December, 2015.)

Patrick, John. 1679. *A Century of Select Psalms, and Portions of the Psalms of David, especially those of Praise. Turned into Metre, and Fitted to the Usual Tunes in Parish Churches.* London: Wing B2536.

Preacher, Kristopher. J. 2001. *Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence.* http://quantpsy.org (11 December, 2016.)

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

Rissanen, Matti. 1999. Syntax. In Roger Lass (ed.), *The Cambridge history of the English language, vol. 3: 1476–1776*, 187–331. Cambridge: Cambridge University Press.

Speyer, Augustin. 2010. *Topicalization and stress clash avoidance in the history of English.* Berlin: De Gruyter Mouton. doi:10.1515/9783110220247

Temperley, Nicholas. 1979. *The music of the English parish church*, 2 vols. Cambridge: Cambridge University Press.

Temperley, Nicholas. 2001. Hymn, § IV: Protestant. In Stanley Sadie (ed.), *The new grove: Dictionary of music and musicians*, vol. 12, 28–35. London: Macmillan.

van der Wurff, Wim & Tony Foster. 1997. Object–verb order in 16th century English: A study of its frequency and status. In Raymond Hickey & Stanisław Pupple (eds.), *Language history and linguistic modelling: A Festschrift for Jacek Fisiak on his 60th birthday*, vol. 1, 439–453. Berlin: Mouton de Gruyter. doi:10.1515/9783110820751.439

Wessa, Patrick. 2015. *Pearson correlation (v1.0.10) in free statistics software (v1.1.23-r7).* Office for Research Development and Education. http://www.wessa.net/rwasp_correlation.wasp/ (11 December, 2016.)

Wither, George. 1624. *The Schollers Purgatory, Discourered In the Stationers Common-wealth, And Discribed in a Discourse Apologeticall, aswell for the publike aduantage of the Church, the State &, whole Common-wealth of England, as for the remedy of priuate iniuryes.* London: STC 25919.

Appendix

Argument shifting in congregational song per author, split up into syntactic functions
(absolute frequencies; *BP* = basic population (all arguments that may be shifted), *AS* = instances
of argument shifting)

Year	Author	No. of words	Argument shifting		Object shifting		Complement shifting		Shifting of obligatory adverbials	
			BP	AS	BP	AS	BP	AS	BP	AS
1535	Coverdale	*2005*	*278*	*68*	*167*	*51*	*45*	*5*	*66*	*12*
1549	Crowley	*2006*	*231*	*52*	*149*	*36*	*35*	*5*	*47*	*11*
1549	Wyatt	*1995*	*187*	*48*	*125*	*31*	*26*	*3*	*36*	*14*
1549	Howard	*1999*	*231*	*86*	*123*	*55*	*37*	*9*	*71*	*22*
1553	Seager	*2009*	*227*	*104*	*156*	*79*	*20*	*2*	*51*	*23*
1562	Sternhold and Hopkins	*2001*	*226*	*70*	*162*	*51*	*32*	*7*	*32*	*12*
1565	Hall	*2002*	*224*	*85*	*134*	*52*	*36*	*8*	*54*	*25*
1567	Parker	*1997*	*263*	*109*	*164*	*67*	*31*	*11*	*68*	*31*
1569	Samuel	*2002*	*259*	*105*	*143*	*64*	*47*	*18*	*69*	*23*
1578	Wedderburn	*1996*	*278*	*88*	*172*	*52*	*39*	*12*	*67*	*24*
1583	Hunnis	*2004*	*198*	*82*	*121*	*51*	*31*	*11*	*46*	*20*
1591	Drayton	*2010*	*211*	*80*	*144*	*52*	*19*	*1*	*48*	*27*
1595	Southwell	*2000*	*211*	*78*	*118*	*44*	*41*	*9*	*52*	*25*
1597	Lok	*2000*	*228*	*116*	*131*	*70*	*41*	*14*	*56*	*32*
1599	Mary S. Herbert	*1997*	*221*	*103*	*146*	*65*	*24*	*9*	*51*	*29*
1601	Verstegan	*2002*	*239*	*112*	*137*	*64*	*39*	*13*	*63*	*35*
1612	Ainsworth	*2008*	*248*	*118*	*112*	*40*	*56*	*36*	*80*	*42*
1620	Dod	*2002*	*216*	*102*	*125*	*57*	*37*	*24*	*54*	*21*
1623	Wither	*1996*	*230*	*120*	*164*	*81*	*30*	*15*	*36*	*24*
1644	Barton	*1995*	*225*	*105*	*142*	*70*	*39*	*14*	*44*	*21*
1651	King	*1997*	*255*	*141*	*163*	*92*	*24*	*6*	*68*	*43*
1668	Smyth	*2000*	*231*	*98*	*154*	*67*	*28*	*8*	*49*	*23*
1668	Birchley	*2001*	*197*	*38*	*131*	*26*	*26*	*3*	*40*	*9*
1679	Patrick	*2000*	*240*	*98*	*148*	*64*	*39*	*9*	*53*	*25*
1682	Reeve	*1998*	*225*	*106*	*128*	*62*	*35*	*8*	*62*	*36*
1691	Keach	*2002*	*247*	*101*	*107*	*46*	*70*	*29*	*70*	*26*
1691	Mason	*1993*	*234*	*60*	*129*	*35*	*37*	*4*	*68*	*21*
1692	Baxter	*2008*	*255*	*111*	*155*	*61*	*41*	*20*	*59*	*30*
1696	Tate and Brady	*2001*	*200*	*97*	*149*	*71*	*19*	*5*	*32*	*21*
1697	Woodf. / Herbert	*2001*	*228*	*64*	*155*	*43*	*30*	*10*	*43*	*11*
1707	Watts	*2002*	*209*	*37*	*134*	*16*	*28*	*6*	*47*	*15*
1741	Cennick	*2000*	*234*	*54*	*173*	*44*	*28*	*5*	*33*	*5*
1745	Hammond	*1998*	*241*	*72*	*179*	*54*	*27*	*3*	*35*	*15*

1755	Doddridge	2004	209	72	156	52	24	4	29	16
1757	Allen and Batty	2001	275	87	185	63	26	5	64	19
1759	Hart	2000	249	59	146	42	56	5	47	12
1759	Toplady	2004	265	67	178	45	36	5	51	17
1765	Merrick	2002	221	119	159	78	5	0	57	41
1769	Gibbons	1999	193	54	119	23	21	3	53	28
1773	Byrom	1997	196	27	126	19	28	1	42	7
1779	Cowper and Newton	1998	233	51	158	37	34	2	41	12
1780	Wesley	2000	216	81	163	65	22	8	31	8
1780	Steele	1997	189	35	130	23	20	1	39	11
1782	Fawcett	2001	227	58	151	38	30	3	46	17
1792	Swain	2003	202	51	148	39	18	2	36	10
1824	Conder	2002	214	44	122	28	48	4	44	12
1827	Montgomery	2001	199	50	132	29	32	5	35	16
1827	Heber	2003	179	51	122	37	13	3	44	11
1827	Keble	1998	176	33	116	21	18	1	42	11
1837	Chandler	2005	184	66	127	45	19	4	38	17
1839	Elliott	1999	229	60	139	33	28	6	62	21
1843	Neale	2001	193	35	120	22	24	4	49	9
1849	Faber	2001	191	36	105	23	46	5	40	8
1854	Waring	2001	197	36	109	19	33	4	55	13
1855	Lynch	2000	242	61	143	32	49	13	50	16
1865	Wordsworth	2001	184	68	109	34	33	12	42	22
1868	Gill	2000	228	79	139	46	30	9	59	24
1874	Thring	2000	190	57	135	45	26	2	29	10
1884	Havergal	1999	190	51	121	35	35	7	34	9
1888	Ellerton	2000	170	52	105	27	12	4	53	21

Texts for comparison: Early and Late Modern English secular poetry
(absolute frequencies; *BP* = basic population (all objects that may be shifted),
OS = instances of object shifting)

Author	Title	Year	No. of words	BP	OS
Thomas More	*A merry jest; 'Pageant verses'*	1516 / undated	*2,000*	141	31
John Skelton	*Colin Clout*	1545	*2,025*	114	21
Henry Howard, Thomas Wyatt, and others	*Songs and Sonnettes (Tottel's Miscellany)*	1557	*2,012*	144	47
Edmund Spenser	*The Shepheardes Calender*	1579	*1,913*	121	42
Edmund Spenser	*The Faerie Queene*	1590	*2,027*	132	60
W. Shakespeare	*Sonnets*	1609	*2,020*	133	31
Benjamin Jonson	*A Celebration of Charis*	1640–1641	*1,958*	136	16

Author	Title	Year	No. of words	BP	OS
Abraham Cowley	*Poems*	1656	*2,004*	*106*	*28*
John Milton	*Paradise Lost*	1667	*2,007*	*111*	*15*
John Dryden	*MacFlecknoe*	1682	*2,015*	*129*	*34*
Alexander Pope	*The Poetical Works of A. Pope*, Vol. I	1720–1730	*1,961*	*127*	*33*
Samuel Johnson	*Poems*	1750–1760	*2,011*	*166*	*30*
William Cowper	*Poems*	1773–1784	*2,073*	*138*	*17*
William Blake	*Songs of Innocence and of Experience*	1789, 1794	*2,013*	*144*	*25*
W. Wordsworth and Samuel T. Coleridge	*Lyrical Ballads*	1798	*2,010*	*104*	*15*
George Gordon Lord Byron	*Hours of Idleness and other Early Poems*	1800–1805	*2,034*	*139*	*30*
Percy B. Shelley	*The Revolt of Islam. A Poem in Twelve Cantos*	1817	*2,025*	*77*	*9*
Alfred Lord Tennyson	*The Early Poems*	ca. 1830	*2,017*	*71*	*5*
Robert Browning	*The Pied Piper of Hamelin. A Child's Story*	1842	*2,092*	*112*	*6*
Thomas Hardy	*Wessex Poems*	1898	*2,005*	*124*	*14*

Lexis, morphology, and a changing society

Common to the North of England and to New England

British English regionalisms in John Russell Bartlett's *Dictionary of Americanisms*

Javier Ruano-García

University of Salamanca

This paper examines the British English dialect material behind the compilation of John Russell Bartlett's *Dictionary of Americanisms*. It focuses on evidence furnished by six historical dialect glossaries and dictionaries quoted by Bartlett, which include John Ray's *A collection of English words not generally used* (1674, 1691) and James O. Halliwell's *Dictionary of archaic and provincial words* (1847). My aim is to determine the function of these works in Bartlett's dictionary, exploring their lexicographical treatment and their geographic labels, so as to ascertain whether the transatlantic link specified between varieties of American and British English relied on any of them in particular. I also examine the four editions of Bartlett's dictionary and study the impact of the policy concerning the admission of Americanisms on the material found in British English regional sources.

Keywords: *Dictionary of Americanisms*, British English dialects, Americanisms, dialect lexicography, Late Modern English

1. Introduction

In 1848 John Russell Bartlett (1820–1889) wrote that he had "examined all the English provincial glossaries, and the principal English dictionaries" for the compilation of the first edition of his renowned *Dictionary of Americanisms*, which he found essential so as to "know what words and phrases were still provincial in England" (1848: xxvii). Bartlett's motivation was to give a detailed account of the words that were used in 19th-century America – that is, words of American origin, as well as colloquialisms "used in familiar conversation" (1848: iv) and words found

in provincial use, since "the dialects and provincialisms of those parts of England … have extended to New York, Ohio, Indiana, Illinois and Michigan" (1848: iii). The first edition of the *Dictionary* made room, as a result, for an important number of regionalisms that were used on the other side of the Atlantic, and that Bartlett identified thanks to the evidence furnished by a variety of written sources on British English dialects. The second edition of 1859 showed, by contrast, a change of policy concerning the inclusion of words used in England, to the extent that almost 800 terms were omitted in favor of a "more strictly American" (1859: v) compilation, as Bartlett pointed out. The neglect of this amount of dictionary material reflects, in the words of Crowell (1972: 240), Bartlett's "new view of his task as a lexicographer and a new definition of the term 'Americanism'." Similarly, the fourth edition of 1877, which built on the second edition of 1859, deleted more terms, most of which were documented in British English sources.

This paper explores Bartlett's reliance on English regional sources for the compilation of the *Dictionary of Americanisms*. I examine the four editions of the dictionary and undertake a quantitative and comparative analysis of the evidence furnished by six of the most relevant historical dialect glossaries and dictionaries quoted by Bartlett: John Ray's *A collection of English words not generally used* (1674, 1691), Francis Grose's *A provincial glossary* (1787), John T. Brockett's *Glossary of North Country words* (1825), Robert Forby's *The vocabulary of East Anglia* (1830), William Holloway's *A general dictionary of provincialisms* (1838), and James O. Halliwell's *Dictionary of archaic and provincial words* (1847). More specifically, I address Bartlett's treatment of these dialect sources to determine, first, the lexicographical purpose with which they were quoted; second, I attempt to ascertain whether Bartlett relied more heavily on any of these works and, therefore, whether the transatlantic link made explicit for some terms was reliant on any of them. At the same time, the change of policy visible in the second and fourth editions is examined in relation to this dialect material to find out how it affected the treatment of British English regionalisms quoted by Bartlett. To address these issues, this paper is structured as follows. First, it provides a description of Bartlett's dictionary, with a focus on the general aims of this work, the sources used for the compilation, and the concept of Americanism which was shaped and redefined from 1859. It then examines the role of British English regional sources in the dictionary by carrying out a quantitative analysis of the material extracted from each of the six works considered, paying attention to their lexicographical treatment and the labels used across the four editions of the dictionary.

2. John Russell Bartlett's *Dictionary of Americanisms*

2.1 General aims and sources

The first edition of John Russell Bartlett's *Dictionary of Americanisms*: *A glossary of words and phrases usually regarded as peculiar to the United States* (*DOA*) was well received, as reviews reveal. *The American Whig Review*, for example, concluded that "it is as complete, perhaps, as could be expected for such a work and may serve as a good purpose in separating the language of coarseness from that of elegance" (Anonymous 1848: 328), while *The North American Review* of 1849 emphasized the pioneering quality and comprehensiveness of Bartlett's project, highlighting that "[s]ince the valuable vocabulary of Mr Pickering, no serious attempt has been made to collect our peculiar forms of speech" (Anonymous 1849: 94). In fact, Bartlett's dictionary was and is still looked upon as an important addition and achievement with regard to its predecessors, "for it shows that our language is alive," as Lowell (1859: 638) pointed out, and because it was "an admirable attempt to break new ground," in the words of Cassidy & Hall (2001: 217).[1] As is known, the *DOA* ran into three editions later in the 19th century – Bartlett (1859, 1860, 1877) – but the third, 1860 edition was basically "a reprint of the second without alterations" (1860: iii).[2] Its popularity was likewise felt overseas, which is evidenced by M. Keijzer's and Friedrich Köhler's Dutch and German translations of 1854 and 1866, respectively. Nowadays Bartlett's dictionary remains an important work in the field of American lexicography, serving the purposes of national scale projects like the *Dictionary of American regional English* (Cassidy & Hall 1985–2013), *DARE* (see Ellis & Montgomery 2011: 341).

The main motivation behind the compilation of the dictionary was to provide a thorough and comprehensive account of the words that were peculiar to the US by the middle of the 19th century. As a matter of fact, its coverage went beyond words of "purely American origin" (*DOA*1: iv). It also included colloquialisms "used in familiar conversation, and but seldom employed in composition" (*DOA*1: iv), along with English provincial words that had spread and were found in American cities and states like New York, Michigan, Illinois, etc. This, as Crowell (1972: 232) puts it,

---

1. Besides Pickering's *Vocabulary* of 1816, earlier specimens of American words include a list of 38 terms in the introduction to the second edition of Boucher's *Glossary of archaic and provincial words* (1832), a short list appended to David Humphreys' *The Yankey in England* (1815), Robley Dunglison's glossary (1931 [1829–1830]), and Mason Peck's glossary of 1834 (Mencken 1937: 35–36). Fisher (2001: 69) states that these early collections of Americanisms "were compiled more for entertainment and social criticism than for practical use."

2. For the sake of clarity, the editions of Bartlett's dictionary are referred to as *DOA*1, *DOA*2, and *DOA*4.

shows that Bartlett's initial interest "was in 'the vulgar language of the United States', not in Americanisms in any strict sense of that term. It was his intention to produce a book which would range over the whole area of American colloquial speech."

The sources that Bartlett used for the *DOA*1 comprise works representative of American and British English. They furnished forms, meanings, and quotations to gloss and illustrate a large body of words included within a compilation that was relatively simple in terms of structure (see Ruano-García et al. 2015: 105–107). As regards the American materials, Bartlett drew upon previous works such as Webster's *American Dictionary* (n.d.) and Pickering's *Vocabulary* (1816), humorous dialect writing, newspapers, and a series of correspondents who provided him with first-hand information on Western and Southern regionalisms. Crowell (1972: 234) explains that the *DOA*1 relied extensively on Webster's and Pickering's compilations. References to Webster, apparently to the 1828 edition of the *American Dictionary*, are included in over 300 of the *DOA*1 entries, while Pickering's *Vocabulary* of 1816 is quoted in at least 130 entries (see Bailey 2009: 281–202). As regards the British materials, Bartlett relied on an important number of British English works that gave written and authoritative illustration for those terms he heard in North America too. Among them, special notice should be given to some examples representative of standard usage such as the dictionaries of "Todd's Johnson, Richardson, Knowles, Perry, Sheridan, Ash, and many others" (Mueller 2006: 33). Also found are Bartlett's references to the literary works of Chaucer, Spenser, Shakespeare; magazines like *The Athenaeum* and *The Spectator*; and Joseph Addison. In addition, Bartlett explained that he looked at "all the English provincial glossaries, and the principal English dictionaries; which it was necessary to do to know what words and phrases were still provincial in England" (1848: xxvi). These documents feature prominently among the British English dictionary sources, especially since they served Bartlett's endeavor to document the transatlantic links for many terms he found in the US:

> On comparing these familiar words with the provincial and colloquial language of the northern counties of England, a most striking resemblance appeared … . In fact, it may be said, without exaggeration, that nine tenths of the colloquial peculiarities of New England are derived directly from Great Britain; and that they are now provincial in those parts from which the early colonists emigrated, or are to be found in the writings of well accredited authors of the period when that emigration took place. Consequently, it is obvious, that we have the best authority for the use of the words referred to.   (*DOA*1: iii–iv)

As a result, the *DOA*1 included an important amount of vocabulary characteristic of New England, which is not surprising given that Bartlett was a native of Providence, Rhode Island.

The *DOA*2, which was published in Boston, relied on the written sources of the first edition and at the same time benefitted from the assistance afforded by friends

from different parts of the country "who took an interest in the subject" (*DOA*2: iv). Bartlett acknowledged their help regarding terms used east or west, as well as concerning terms denoting American trees, plants, and fruits. These constituted some of the additions to the *DOA*2, which also included an important number of words used in the border areas with Mexico that had not been recorded in the *DOA*1. The edition of 1859 recorded, in fact, words peculiar to the US in far greater numbers, "probably twice as many as the first edition" (*DOA*2: iii).

Although the *DOA*4 likewise built on its predecessors, the events that took place in the years from 1859 and 1877 determined the choice of information sources and the terms recorded. The Civil War, of course, had an impact on this last edition, which "included quite a number of Americanisms arising from that conflict" (Cassidy & Hall 2001: 201). Examples such as *bummer* "An idle, worthless fellow without any visible means of support" and *ku-klux* "Originally a secret organization in some of the Southern States, but which subsequently … resorted to murder to carry out their purposes" are illustrated with quotes from newspapers like *The New York Herald* and civil war songs like *General Boom of the C.S.A.*, respectively.[3] In line with previous editions, and testifying to the changes that America and American English had undergone, the *DOA*4 also relied on humorous literary writings such as Mark Twain's *Roughing it* (1872), which provided mining terms found in Nevada, for example (*DOA*4: vi–vii).

While the publication of the *DOA*1 laid the foundation for subsequent editions and became "the classic example among books of this kind" (Krapp 1925: 376), the type of sources and terms glossed later in 1859 and 1877 varied not only because of the dictionary additions. Actually, a change of policy was implemented that was largely dependent on Bartlett's understanding of the term 'Americanism.'

2.2 Americanisms

As I have already pointed out, the publication of the *DOA*1 did not go unnoticed and was reviewed positively. Crowell (1972: 236) mentions that "it is hardly surprising that the book was welcomed in the pages of American reviews. There was criticism, to be sure, but it was usually accompanied by considerable praise." Much of this criticism was directed at the inclusion of terms that were "more properly English than American" (Lowell 1859: 639). Felix Flügel (1820–1904), a German philologist, called for "a sifting of material," given that Bartlett had not "kept himself more strictly to his task of supplying a dictionary of Americanisms, and [gave] many words and expressions which have exactly the same acceptation in England

---

3. Bartlett notes that *bummer* was "A word much used by our soldiers during the late civil war" (*DOA*4: s.v.).

as in America" (quoted in Crowell 1972: 237). The same view was held by Richard Grant White (1822–1885), one of Bartlett's fiercest critics, who attacked his project in eight long reviews published in *The Atlantic Monthly* between 1878 and 1879, in an attempt to discuss and define the scope of the term 'Americanism'.[4] In the earliest of these reviews, White condemned the dictionary on the ground that it was filled with

> words and phrases which are English pure and simple, English by origin, English by continued usage from time immemorial to the present day, either in colloquial use or in literature, or in both, and which in fact lack nothing required for the completeness of their Englishhood. The effect of such a publication is one of gross and injurious misrepresentation. (1878: 497)[5]

Although White's criticism came after the publication of the *DOA*4, Bartlett had reacted to such critical views since the first edition was published, trying to overcome and improve the deficiencies that different reviews had pointed out.

During the decade between the publication of the *DOA*1 and the *DOA*2, Bartlett rearranged his dictionary materials, producing a "more strictly American" (*DOA*2) compilation that "greatly better[ed] its forerunner" (Lowell 1859: 639). Improvements were clear, first, in the criteria by which Americanisms were admitted into the dictionary. Bartlett specified nine different classes of words that were subsumed under the term 'Americanism': archaisms, English words used with a different meaning in the US, words whose original meaning had been retained in the US but not in England, English provincialisms that were in general use in the US, words recently coined, borrowings from European languages such as Dutch and Spanish, 'Indian' words, 'Negroisms', and terms showing differences in pronunciation (1859: viii). Although this classification was not included in the introduction to the *DOA*1, it seems that Bartlett's understanding of the term 'Americanism' across the two editions was essentially the same.[6] As a matter of fact, the *DOA*1 includes words like *canticoy* "An Indian word, denoting a dancing assembly, still used by aged

---

4. Crowell (1967: 126) argues that White's aim, in his insistence on denying the existence of American English as a variety different from British English, was "to cut the 'huge dictionary' down to size."

5. In line with this, *The Nation* published on 11 March 1897 a review on the occasion of the publication of Part 2 of the *English dialect dictionary* that commented on Joseph Wright's treatment of Americanisms, and remarked that "it is but right to warn British (and Scottish) lexicographers against the notion that there is any such thing in existence as a real dictionary of real Americanisms" (Anonymous 1897: 13).

6. Despite the criticism received during the second half of the 19th century, it is worth highlighting that "Bartlett's criteria were much like our own, admitting Americanisms on both diachronic and synchronic grounds," as Cassidy & Hall (2001: 200) explain with regard to *DARE*.

people in New York and on Long Island", terms showing errors of pronunciations (e.g. *cupalo* for *cupola*), and words that he thought had fallen into disuse by 1848, such as *buss* "a kiss".[7] Interestingly, Bartlett's acceptance of English provincialisms within the category of Americanisms was echoed by later collections and essays on the topic. Examples are William C. Fowler's classification in *The English language* (1850), in which he lists "old words and phrases which are now merely provincial in England," and Sylva Clapin's *New dictionary of Americanisms* (1902), which includes "genuine English words, obsolete or provincial in England, and universally used in the United States" (Mencken 1937: 98–101).

Further refinements of the *DOA2* were visible in Bartlett's policy concerning the exclusion of words, which was of course dependent on the criteria underlying the admission of terms into the dictionary. He rejected almost 800 words that were "common to the colloquial language of England and this country" (1859: iv); examples are *afeard* "afraid" and *to bamboozle* "to deceive", which had been quoted from Todd's edition of Johnson's *Dictionary* (1818). It is worth noting, however, that the *DOA2* did not reject all the words used on both sides of the Atlantic that represented no distinctive American usage. Bartlett retained many of them on account of the fact that they "were now used only in some out-of-the-way locality in England, but quite general here, [and] may be regarded as a peculiarity of the English language as spoken in America, i.e. an Americanism" (1859: iv–v). He was referring to many regional terms that were found in England and recorded in provincial glossaries and dictionaries. It seems that these words were more resistant to editorial intervention; some terms used in provincial speech were even added to the *DOA2*, as I show in the following section.

The *DOA4* continued with Bartlett's endeavor to produce a genuinely American compilation. There were further deletions, most of which affected British English words added to the *DOA2*; according to Crowell (1972: 241), they represented "little more than a sharpening of the policy he had set earlier." An example is *cobble-stone* "a roundish stone" of which he had noted in the *DOA2* that "Mr. Halliwell informs us that the word *cobble* is used in this sense in the North of England." Like *cobble-stone*, other terms quoted from English dialect sources were removed from the *DOA4*, which raises questions concerning (i) how the policy of word exclusion applied to British English regionalisms and the sources by which they were represented and (ii) the share of dictionary material represented by these words that found a place within the category of Americanism.

---

7. *DARE* shows that *buss* was, in fact, still in use during the 20th century. It records 20th-century examples from journals such as *Dialect Notes* (1927) and *American Speech* (1935), as well as the evidence furnished by informants from New Mexico, Michigan, Pennsylvania, or Florida interviewed between 1965 and 1970. I thank an anonymous reviewer for kindly bringing this to my attention.

3.   British English regionalisms in the *Dictionary of Americanisms*

As indicated above, provincial documents feature prominently among the British English dictionary sources used for the compilation of the *DOA*. Bartlett consulted a series of regional works that comprise literary material such as Anderson's *Ballads in the Cumberland dialect* (1808), the anonymous Essex poem *John Noakes & Mary Styles* (1839), and Hamilton's *Nugae Literariae* (1841), as well as non-literary documents: glossaries of localized varieties such as Carr's glossary of the Craven dialect (1828) and Hartshorne's *Salopia Antiqua* (1841). In addition, Bartlett quoted from more general dialect works that deserve special attention for their importance in the historical record of English regional dialects: Ray (1674, 1691), Grose's *A provincial glossary* (1787), Brockett's *Glossary of North Country words* (1825), Forby's *The vocabulary of East Anglia* (1830), Holloway's *A general dictionary of provincialisms* (1838), and Halliwell's *Dictionary of archaic and provincial words* (1847).

The main lexicographical function of these sources was to provide a definition for the words recorded: Figure 1 shows that over 50% of the material extracted from these dialect works was employed as definitions in the *DOA*1 (50.86%), *DOA*2 (51.85%), and *DOA*4 (60%).
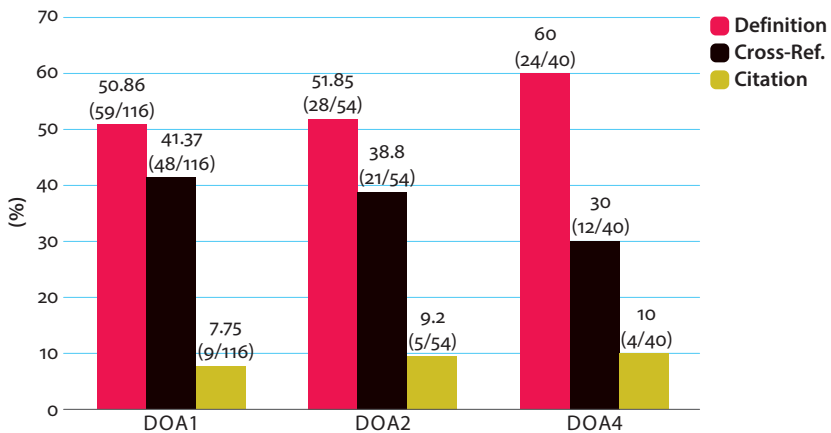
Figure 1. Lexicographical function of British English regional sources in the *DOA*: percentages[8]

These sources were likewise used cross-referentially to indicate that further information on specific senses or forms could be found in some of them. An example is the entry for *chunk* "a short, thick piece of wood" in which Ray's and Grose's

---

8.   The data represented in Figure 1 does not include the 19 words that were added to the *DOA*2 (Table 1), on which the *DOA*4 built. As such, the material in the *DOA*2, as represented in Figure 1, comprises 54/73 items, while that of the *DOA*4 comprises 40/59 items (Table 2).

glossaries are referenced for information on the southern English distribution of this noun. A smaller number of entries were also employed as citation material in the *DOA* too, as in:

(1) PEAKED. Thin and emaciated, as from sickness. Holloway
     says, that in England they say of a sickly person, "he
     looks pale and *peaked*." The same expression is often heard
     in the Northern States.   (*DOA*2 s.v.)

It is worth underlining that not all these sources, if examined individually, played the same role in the *DOA*. Ray (1674, 1691) and Grose (1787), for example, were mostly used as cross-references in the *DOA*1, in 3/5 and 11/15 instances, respectively, e.g. *hunk* "a large piece or slice; a big lump" and *swale* "to melt and run down, as the tallow of a candle". Something similar applies to the words quoted from Holloway (1838) in the *DOA*1, as this dictionary was more often employed as citation (in 4/15 instances) than as cross-reference material (in 1/15 instances). The global results displayed in Figure 1 do not mirror, therefore, the specific treatment that Bartlett gave to each of these works. This is basically due to the fact that the *DOA* did not quote from them on the same basis, and that Bartlett made use of Halliwell (1847) and Forby (1830) more frequently.

As a matter of fact, references to Forby (1830) and Halliwell (1847) predominate over the rest of the sources in the *DOA*1. Table 1 shows that, together, they were quoted in 56.8% of the entries (66/116) where Bartlett acknowledged the assistance afforded by the works examined. This is also seen in the *DOA*2 in which both of them were referenced more extensively: Halliwell (1847) was quoted on 25 occasions, followed by Forby (1830) with 19 quotes.

Table 1. British English regional sources in the *DOA*1 and *DOA*2: raw figures[9]

	*DOA*1	*DOA*2			
		Items deleted	Reference retained	Reference removed	New to *DOA*2
Ray	5	2	1	2	0
Grose	15	8	5	2	1
Brockett	15	9	2	4	3
Forby	33	18	10	5	4
Holloway	15	6	6	3	0
Halliwell	33	19	9	5	11
Subtotal		*62*	*33*	*21*	*19*
Total	116			73	

---

9. The total of 73 items identified in the *DOA*2 comprises the references retained, the references removed, and the words added to the second edition of the dictionary.

The differences between both editions can first be considered quantitatively. Visibly, they are in line with the change of policy that Bartlett implemented from 1859. The figures in Table 1 show, for example, that 53.3% of the words illustrated by Grose (1787) in the *DOA*1 (8/15) were deleted from the *DOA*2 – *chink* "a term for money", *slab* "the outside of a piece of timber or log" – and that 54.5% of the material (18/33) that Bartlett quoted from Forby (1830) was not admitted into the second edition, e.g. *bobbery* "a squabble, a row" and *crimany* "interj. of sudden surprise". Secondly, differences lie in the fact that some of the references to these regional sources were removed from the corresponding entries in the *DOA*2, as in:

(2) a. *DOA*1
         SIGHT. A great many. – *Brockett, Glossary*. … *Sight* is used in most of the Northern and Eastern, and *heap* in the Southern and Western States.
     b. *DOA*2
         SIGHT. A great many; a great deal. … *Sight* is used in most of the Northern and Eastern, and *heap* in the Southern and Western States.

There is no clear explanation in the introduction to the *DOA*2 that may account for this type of omission. One could assume that it was perhaps an editorial decision of Bartlett in compliance with his aim of making the *DOA*2 a 'more strictly American compilation', and that, as a result, he veiled the links with England in the entries for those words that were current on both sides of the Atlantic but did not fall into the category of Americanism. As (2) reads, *sight* was provincial in England but in general use in the US, which was one of the criteria by which words were admitted into the *DOA*2.[10] Omitting information regarding its British English distribution may have thus served to justify its treatment as a proper Americanism, with a gloss that gave *sight* as characteristic of Northern and Eastern America.

Thirdly, and despite the omissions detected, the second edition included additional terms recorded in these regional sources. As illustrated in Table 1, they are relatively few (19 terms) in comparison with those that were deleted (62 in total). Examples are *brewis* "a pottage made of slices of bread with fat broth poured over them" in Halliwell (1847), *to snort* "to laugh outright" in Brockett (1825), and *whoosh* "a term used in backing a horse" in Forby (1830).[11]

---

10. Brockett (1825) defined *sight* as "a great deal, many" (s.v. *site, seet*). Note that the form *sight* found in the *DOA*1 was included in the second edition of Brockett's glossary (1829), the entry for which reads: "vulgarly pronounced *sâet*, a great many" (s.v. *site*, or *sight*).

11. Bartlett remarked that "I've never heard this word except in Nantucket," and quoted it from Suffolk too in light of Moor's Suffolk glossary where "it is defined as 'an imperative, commanding the fore-horse of a team to bear to the left'" (*DOA*2 s.v. *whoosh*).

The same type of changes and differences are found in the *DOA*4. Table 2 shows, on the one hand, that the policy of word exclusion affected more terms found in English regional sources. It had an effect on all the sources examined, with the exception of Ray (1674, 1691), and was applied both to the terms that had remained intact since 1848 (e.g. *well to do* "in a state of ease as to pecuniary circumstances", found in Holloway 1838), as well as to those that had been added to the *DOA*2 (e.g. *larrup* "to beat, to flog"; quoted from Forby 1830).

Table 2. British English regional sources in the *DOA*4: raw figures

		Ray	Grose	Brockett	Forby	Holloway	Halliwell
*DOA*4	Reference Retained	1	5	4	12	5	13
	Reference Removed	2	2	4	4	3	4
Subtotal		*3*	*7*	*8*	*16*	*8*	*17*
Total				59			

Such neglect of dialect material is clearly seen in the rejection of 32% of the words (8/25) that Bartlett had quoted from Halliwell (1847) in the previous edition, e.g. *taffy* "a kind of candy made of molasses, flour, and butter, baked in a pan" and *tarnation* "a common oath". His stronger reliance on Halliwell (1847) is evident in the *DOA*4, though.

On the other hand, and in line with what is observed in the *DOA*2, Bartlett omitted quotes and references to dialect sources from some of the dictionary entries (19/59, 32.2% in total). An example is *cow-lease*:

(3)   a.   *DOA*2
          COW-LEASE. A right of pasturage for a cow in a common pasture. Used in some towns of New England. –
          Pickering's Vocabulary. Provincial in the West of England. – *Grose's Glossary*.
     b.   *DOA*4
          COW-LEASE. A right of pasturage for a cow in a common pasture. New England. – *Pickering*. Provincial in the west of England.

Likewise apparent are the quantitative differences concerning the geographic labels deployed by Bartlett (see Table 3). Regarding the *DOA*1, a large amount of the material quoted from the regional sources (80/116 words) was characterized by Bartlett as widespread American use, often with comments like "with us it is frequently heard," as in the entry for *gumption* "understanding". Table 3 also shows that 23.3% of the entries (27/116) were ascribed to New England (e.g. *to squiggle* "to move about like an eel") while the rest of the data (7.75%) were localized in other states, such as Louisiana (e.g. *honey-fogle* "to swindle"), or marked with stylistic or social comments (e.g. *honor bright* "a protestation of honor among the vulgar").

Bartlett's stronger reliance on Forby (1830) and Halliwell (1847) may have been due to the fact that they furnished the highest number of regional attestations for items he found to be in general American distribution. As Table 3 furthermore suggests, Forby and Halliwell included a relatively important number of cases that exemplified the British link for words used in Bartlett's native New England.

Table 3. American distribution of the words attested in English regional sources: *DOA*1

	Ray	Grose	Brockett	Forby	Holloway	Haliwell	*Subtotal*
US general	4	9	10	23	10	24	*80*
New England	1	5	3	9	4	5	*27*
Others	0	1	2	1	1	4	9
Total	5	15	15	33	15	33	116

In this last respect, it is worth looking into the English distribution of the New England words quoted. As displayed in Figure 2, the majority of them are found in the North of England and East Anglia, with 15 instances of the sample (55.5%). This reflects in some way the original patterns of settlement visible in the waves of East Anglian and Northern English migrants coming to New England from the early 17th to the late 18th centuries (see Algeo 2001: 7–8; Fisher 2001: 60). Actually, Craigie argues that "in New England three areas are chiefly represented – Yorkshire and Lincolnshire, East Anglia and the Southwestern counties" (quoted in Mencken 1937: 129).[12] It seems no coincidence therefore that the *DOA*1 consulted Forby (1830) on a regular basis; actually, all the New England words linked to East Anglia were illustrated by means of this glossary. Similarly, it seems that the lexicographical impact of Halliwell (1847), and, obviously, of Brockett (1825), was related to the Northern English ascriptions provided.

Concerning the *DOA*2, the data are in line with the above results. A larger number of words are reported as general Americanisms (39/73; 53.4%), followed by terms ascribed to New England (26/73; 35.61%) and examples characteristic of other areas like the South West (8/73; 10.9%), e.g. *safe* "sure, certain". The analysis interestingly indicates that the New England items were more resistant to the policy by which Bartlett regulated the admission of words into the *DOA*2. As a matter of fact, 96.29% of the vocabulary ascribed to New England in the *DOA*1 was retained in the second edition (26/27 words), while only 48.75% of the general Americanisms found in the first edition was preserved in the *DOA*2 (39/80 words). Most of the New England additions can also be traced to the North and the East

---

12. Note that Robert Willan's list of words of the West Riding of Yorkshire (1814) and the glossary of the Craven dialect (1828) are among the regional sources often quoted by Bartlett, especially in the first edition.
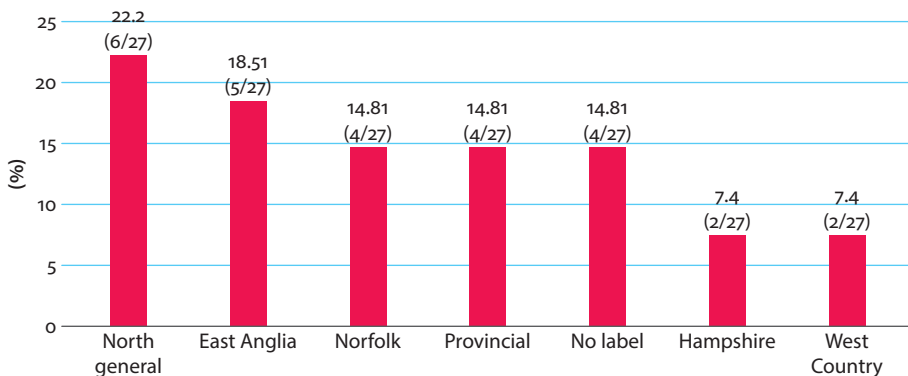
Figure 2. English distribution of the words Bartlett ascribed to New England (*DOA*1): percentage

of England, with Forby (1830) and Halliwell (1847) furnishing most of the written evidence that Bartlett quoted.

The same pattern of distribution is observed in the *DOA*4. Unlike in previous editions, however, the quantitative difference between the dictionary data representative of general Americanisms and of New England items is less prominent: they represent 44.06% (26/59) and 42.37% (25/59) of the material examined, respectively. This indicates that Bartlett omitted a larger portion of words that he thought were used in widespread American distribution; in fact, the *DOA*4 rejected 33.3% (13/39) of the general Americanisms counted in the *DOA*2 (e.g. *teeter-tawter* "the game of see-saw"), whilst it excluded only 3.8% (1/26) of the vocabulary ascribed to New England (e.g. *beaker* "a large drinking-glass, a tumbler"). The results suggest that these words escaped the excisions made to the *DOA*4 more frequently, probably because they were understood as genuine Americanisms that, despite their use on the other side of the Atlantic, Bartlett knew from first-hand experience as a native of the northeast of the US.

4. Concluding remarks

This paper has addressed the role of British English regionalisms in John R. Bartlett's renowned *Dictionary of Americanisms*. I have explored the function of six important British English historical dialect works cited by Bartlett, with the aim of determining the extent to which his policy of admission of Americanisms had an effect on the British English regionalisms quoted and the sources by which they were represented, as well as whether the links specified between varieties of

British and American English were reliant on any of these sources. The comparative study of the four editions of the dictionary has highlighted a quantitative change regarding the data taken from the six sources examined, with a decrease by 49% of the number of British English regionalisms quoted: the *DOA*1 has 116 instances, whilst the *DOA*4 has 59. This responds to Bartlett's decision to reject much of the dictionary material that apparently represented no distinctive American usage, as contemporary critics had suggested. In fact, the analysis has shown that excisions affected all six works under investigation, though the words listed by Forby (1830) and Halliwell (1847) appear to have resisted editorial cuts more often. This may be in line with Bartlett's heavy reliance on both of them and with the fact that some of the material extracted from these two sources was among the additions included in the edition of 1859. It is likely that the important contribution of these two works lies in the fact that both Forby (1830) and Halliwell (1847) furnished the highest number of attestations for provincial words that were at the same time widespread in American use, one of the criteria by which Bartlett admitted Americanisms into his compilation.

The study has likewise suggested that Bartlett's lexicographical coverage of American English drew on these works more extensively for words that more frequently occur in East Anglia, Norfolk, and the North of England, which Bartlett linked to, and identified as characteristic of, his native New England. In line with this, quantification of the data reveals that New England terms linked to varieties of East Anglia and Northern English were among those retained in the dictionary on a more frequent basis.

All in all, the evaluation of British English regional data in Bartlett's groundbreaking compilation goes some way towards illuminating how transatlantic varieties of English were understood, related, and represented, and at the same time how these lexicographical links testified to the intense contemporary debate on the concept of Americanism, especially since they reflected the shift of some originally English dialectalisms into words that were seen and accepted as American (see also Minkova, this volume, on a similar shift in register that clipped words have experienced in the history of English). Further investigation into the reception and treatment of other British (non-)dialect works quoted by Bartlett may provide more refined insights into the link between Englishes on both sides of the Atlantic and on the 19th-century idea of Americanism more generally.

References

Primary sources

Anderson, Robert. 1808. *Ballads in the Cumberland dialect*. Wigton: R. Hetherton.

Anonymous. 1839. *John Noakes & Mary Styles; or, "An Essex calf's" visit to tiptree races: A poem*. London: John Russell Smith.

Bartlett, John R. 1848. *Dictionary of Americanisms: A glossary of words and phrases usually regarded as peculiar to the United States*. New York: Bartlett & Welford.

Bartlett, John R. 1859. *Dictionary of Americanisms: A glossary of words and phrases usually regarded as peculiar to the United States*, 2nd edn. Greatly improved and enlarged. Boston: Little, Brown & Company.

Bartlett, John R. 1860. *Dictionary of Americanisms: A glossary of words and phrases usually regarded as peculiar to the United States*, 3rd edn. Greatly improved and enlarged. Boston: Little, Brown & Company.

Bartlett, John R. 1877. *Dictionary of Americanisms: A glossary of words and phrases usually regarded as peculiar to the United States*, 4th edn. Boston: Little, Brown & Company.

Boucher, Jonathan. 1832. *Boucher's glossary of archaic and provincial words: A supplement to the dictionaries of the English language, particularly those of Dr. Johnson and Dr. Webster … Edited jointly by the Rev. Joseph Hunter and Joseph Stevenson*. London: Black, Young & Young.

Brockett, John T. 1825. *A glossary of North Country words, in use: From an original manuscript in the library of John George Lambton, Esq., M.P. with considerable additions*. Newcastle upon Tyne: Printed by T. & J. Hodgson, for E. Charney.

Brockett, John T. 1829. *A glossary of North country words, in use; with their etymology, and affinity to other languages; and occasional notices of local customs and popular superstitions*. Newcastle upon Tyne: Emerson Charnley.

Carr, William. 1828. *The dialect of Craven, in the West-Riding of the county of York, with a copious glossary, illustrated by authorities from ancient English and Scottish writers, and exemplified by two familiar dialogues*, 2 vols. London: Printed for WM. Crofts.

Cassidy, Frederic G. & Joan H. Hall (eds.). 1985–2013. *Dictionary of American regional English*. Cambridge, MA: Belknap Press of Harvard University Press.

Clapin, Sylva. 1902. *A new dictionary of Americanisms; being a glossary of words supposed to be peculiar to the United States and the dominion of Canada*. New York: L. Weiss & Co.

Dunglison, Robley. 1931 [1829–1830]. Americanisms from Virginia literary museum and journal of belles lettres, arts &c. In Mitford M. Mathews (ed.), *The beginnings of American English: Texts and essays*, 99–112. Chicago: The University of Chicago Press.

Forby, Robert. 1830. *The vocabulary of East Anglia; an attempt to record the vulgar tongue of the twin sister counties Norfolk and Suffolk, as it existed in the last twenty years of the eighteenth century, and still exists*, 2 vols. London: Printed by and for J.B. Nichols & Son.

Grose, Francis. 1787. *A provincial glossary: With a collection of local proverbs, and popular superstitions*. London: Printed for S. Hooper.

Halliwell, James O. 1847. *A dictionary of archaic and provincial words*. London: John Russell Smith.

Hamilton, Richard Winter. 1841. *Nugae literariae: Prose and verse*. London: Hamilton, Adams & Co.

Hartshorne, Charles Henry. 1841. *Salopia Antiqua, or an enquiry from personal survey into the druidical, military, and other early remains in Shropshire and the North Welsh borders*. London: J. W. Parker.

Holloway, William. 1838. *A general dictionary of provincialisms, written with a view to rescue from oblivion the fast fading relics of by-gone days*. Lewes: Sussex Press.

Humphreys, David. 1815. *The Yankey in England, a drama in five acts*. [Boston, MA]: [s.n.]

Pickering, John. 1816. *A vocabulary, or collection of words and phrases which have been supposed to be peculiar to the United States of America. To which is prefixed an essay on the present state of the English Language in the United States*. Boston: Cummings & Hilliard.

Ray, John. 1674. *A collection of English words not generally used*. London: H. Bruges.

Ray, John. 1691. *A collection of English words not generally used*. London: Christopher Wilkinson.

Todd, Henry John (ed.). 1818. *A dictionary of the English language, by Samuel Johnson*. London: Longman, Hurst, Rees, Orme & Brown.

Twain, Mark. 1872. *Roughing it*. Hartford: American Publishing Co.

Webster, Noah. N.d. *An American dictionary of the English language*, 2 vols. New York: S. Converse.

Willan, Robert. 1814. A list of ancient words at present used in the mountainous district of the West Riding of Yorkshire. *Archaeologia* 17. 138–167.

Secondary sources

Algeo, John. 2001. External history. In John Algeo (ed.), *The Cambridge history of the English language, vol. 6: English in North America*, 1–58. Cambridge: Cambridge University Press. doi:10.1017/CHOL9780521264792.002

Anonymous. 1848. Critical notices: Dictionary of Americanisms: A glossary of words and phrases, usually regarded as peculiar to the United States, by John Russell Bartlett. *The American Whig Review* 8(3). 326–328.

Anonymous. 1849. Review of Dictionary of Americanisms: A glossary of words and phrases, usually regarded as peculiar to the United States, by John Russell Bartlett. *The North American Review* 69(144). 94–110.

Anonymous. 1897. Review of the English dialect dictionary, part 2, December 1896. *The Nation*, 11 March. Bodleian Library, MS Eng. lang. c. 13: 12–13.

Bailey, Richard. 2009. National and regional dictionaries of English. In Anthony P. Cowie (ed.), *The Oxford history of English lexicography, vol. 1: General-purpose dictionaries*, 279–390. Oxford: Oxford University Press.

Cassidy, Frederic G. & Joan H. Hall. 2001. Americanisms. In John Algeo (ed.), *The Cambridge history of the English language, vol. 6: English in North America*, 184–218. Cambridge: Cambridge University Press. doi:10.1017/CHOL9780521264792.006

Crowell, Michael G. 1967. Richard Grant White and Americanisms. *American Speech* 42(2). 122–130. doi:10.2307/453963

Crowell, Michael G. 1972. John Russell Bartlett's Dictionary of Americanisms. *American Quarterly* 24(2). 228–242. doi:10.2307/2712072

Ellis, Michael & Michael Montgomery. 2011. About all: Studies in nineteenth-century American English. *American Speech* 86(3). 340–354. doi:10.1215/00031283-1503928

Fisher, John Hurt. 2001. British and American, continuity and divergence. In John Algeo (ed.), *The Cambridge history of the English language, vol. 6: English in North America*, 59–85. Cambridge: Cambridge University Press. doi:10.1017/CHOL9780521264792.003

Fowler, William C. 1850. American dialects. In *The English language in its elements and forms. With a history of its origin and development. Designed for use in colleges and schools*. 91–94. New York: Harper & Brothers.

Krapp, George P. 1925. *The English language in America*, vol. 1. New York: The Century Company.

Lowell, James R. 1859. Reviews and literary notices. *The Atlantic Monthly* 4(25). 638–645.

Mencken, Henry L. 1937. *The American language: An inquiry into the development of English in the United States*. New York: Alfred A. Knopf.

Mueller, Jerry E. (ed.). 2006. *Autobiography of John Russell Bartlett (1805–1886)*. Providence, RI: The John Carter Brown Library.

Ruano-García, Javier, María F. García-Bermejo Giner & Pilar Sánchez-García. 2015. "Provincial in England but in common use with us": John R. Bartlett's *Dictionary of Americanisms* and the *English dialect dictionary*. In Marina Dossena (ed.), *Transatlantic perspectives on Late Modern English*, 99–116. Amsterdam: John Benjamins.

White, Richard G. 1878. Americanisms. *The Atlantic Monthly* 41(246). 495–502.

Betwixt, amongst, and *amidst*

The diachronic development of function words with final /st/

Ryuichi Hotta
Keio University

The purpose of the paper is to examine, using historical corpora, the diachronic development of variants of the function words BETWEEN, AMONG, and AMID with emphasis on the variants with final /st/, i.e. *betwixt*, *amongst*, and *amidst*. In Present-day English, the variants with final /st/ have a more formal, literary, or archaic ring to them than their counterparts without it. In older English, however, the former variants were more widely used than they are today. This paper addresses how individual variants – especially ones with final /st/ – came into being, how they competed with one another in terms of frequency in each period, and how this resulted in the Present-day English distribution.

Keywords: Middle English, Modern English, preposition, paragoge, corpus

1. Introduction

This paper examines the diachronic development of variants of the function words BETWEEN, AMONG, and AMID with special emphasis on the variants with final /st/, i.e. *betwixt*, *amongst*, and *amidst*. In Present-day English (PDE), variants with final /st/ have a more formal, literary, or archaic ring than their counterparts without it. In older English, however, the variants with final /st/ were more widely used than they are today. To my knowledge, no diachronic studies have been conducted about the ebb and flow of the variants, and questions remain unanswered as to how individual variants – especially the ones with final /st/ – emerged, how they competed in each period, and how this resulted in the current PDE distribution.

I cover all the historical periods of English from Old English (OE) to PDE, but I focus on the Middle English (ME) and Modern English (ModE) periods, as these are the periods when the several variants occurred most frequently and competition between them was the greatest. To collect as many examples as possible from historical sources, the study draws on historical corpora and text databases as well

as commonly cited dictionaries. In the first place, I consulted the *Oxford English dictionary* (*OED*) and *Middle English dictionary* (*MED*) to appreciate the range of variants historically attested. Then, I searched *The Helsinki corpus of English texts* (*HC*) for variants to get an overview of their distribution in each period and their changing distribution across periods. The overview thus obtained served as a useful starting point, but since *HC* is a small-sized corpus by today's standards, the data obtained from *HC* had to be supplemented with data from historical corpora focusing on particular periods of the language. I used *A linguistic atlas of Early Middle English* (*LAEME*) for Early Middle English (EME), *The Middle English grammar corpus* (*MEGC*) for Late Middle English (LME), *Early English books online* (*EEBO*) for Early Modern English (EModE), *The corpus of Late Modern English texts* (*CLMET3.0*) for Late Modern English (LModE), and *BNCweb CQP-edition* (*BNCweb*) for PDE.

 After providing a historical description of the variants of the three lexical items in the following three sections, I examine several views that have been proposed about the emergence, growth, and decline of the variants with final /st/, and then I attempt to integrate the different views into a likely scenario accounting for the historical changes. Before I go on to the study, some preliminary notes are in order. To be able to refer to the wide range of variants accurately, I have adopted a notational system to facilitate reference at distinct levels. To refer to the lexical items at their most abstract level, I use BETWEEN, AMONG, and AMID, in upper case.[1] These abstract lexical items subsume the variant spellings (written in italics), such as *betux*, *bituhen*, *betweonan* for BETWEEN; *amange*, *emonges*, *on gemang* for AMONG; and *amidde*, *amyddes*, *amid'st* for AMID. The variant spellings can then be classified into a number of types according to their morphological configuration, particularly of their final segments. Within the item BETWEEN, for example, several types can be distinguished by their ending, such as 'betwix', 'betuh', 'betweenen' (these types are indicated by single quotation marks); types are often abbreviated to 'x', 'h', 'nn', and the like. The classification into types is only sensitive to consonantal variation and the final vowel at the end of the spelling variant. In other words, *atwix*, *betweox*, and *be twyx*, for example, all belong to the 'x'-type; *among*, *bimong*, and *on gemang* are members of the 'g'-type; *amidde*, *in mydde*, and *on midde* are assigned to the 'de'-type. Admittedly, the difference in prefixes and vowels is also important enough to deserve independent examination, but here, I am particularly interested in the development of paragogic consonants, and I therefore concentrate on word-final segments.[2]

---

1.   In the present study, I use 'lexical item' in a specific sense. In everyday usage, *between* and *betwix* are referred to as different words or lexical items, but in this study I treat them as alternative realizations of the more abstract entry BETWEEN, which I refer to as a 'lexical item'.

2.   To give an idea of the importance attached to parts of the word that I must leave unexamined, consider that *eLALME* distinguishes sixteen types according to several criteria, giving

Note, finally, that in this study, the different syntactic functions of between, among, and amid (e.g. their part-of-speech function) are not distinguished because the primary focus of this study is morphonological rather than syntactic – though I acknowledge that some items have different meanings depending on their syntactic function.[3]

2. between

2.1 Etymological notes and historical overview based on the *HC*

Historical variants of between abound. A search of the *HC* yielded 101 distinct variants, and searches of other historical corpora yielded dozens more. The variants can be classified, however, into a manageable number of types according to their endings, which largely reflect their etymological makeup and later morphonological processes. According to Kitson (1993: 12), underlying all known variants of between are the component prototypes corresponding to the preposition by and the numeral two. In early OE, at least in poetry, the two elements occurred separately, as in *be sǣm tweonum* "between two seas" (*Beowulf* 858). Later, the transposition of the noun and the numeral (i.e. *be twēonum sǣm*) triggered a syntactic reanalysis that allowed *be twēonum* to be reinterpreted as a composite preposition.

The numeral two inflected in various ways in and before OE; accordingly, a range of forms with different inflections can be attested in the corpora. Three main types can be recognized. Firstly, the 'betwēonum'-type, or 'nm' in abbreviation, is characterized by the dative ending *-um*, as required by the preposition. In late OE, the *um*-ending was phonetically leveled to *en*, thus resulting in the 'betwēonen'-type ('nn'). The second nasal as well as the schwa preceding it, however, were subject to further phonetic attrition, producing the 'betweene'-type ('ne') and eventually the 'between'-type ('n').[4]

---

corresponding dot maps under the item number 89. See also the dot maps 703–706 and 1118–1119 in *LALME*, and Ciszek-Kiliszewska's (2013) separate treatment of *twix* as distinct from *bitwix*.

3. In her study on the preposition-cum-adverb *twix*, Ciszek-Kiliszewska (2013: 88) discusses the semantic differences between the prepositional and adverbial uses, with the former meaning "among, in among; between" and the latter meaning "in the meantime, from this moment".

4. It is likely that, in many cases, the final <e> was a silent <e>, i.e. an orthographical variant without a phonetic equivalent. Thus I could have merged the 'n'- and 'ne'-types, for example, into an 'n(e)'-type. I did not do this, however, for consistency with leveling and loss of inflection: I assumed that the distinction between the 'n'- and 'ne'-types was as important as that between the 'nm'- and 'nn'-types or between the 'nn'- and 'n'-types.

Secondly, the 'betwēox'-type ('x') is characterized by its final *x*. The second component of OE *betwēox* has its origin in Germanic *twa (two) + *-iskaz (-ish); clearly, by the OE period, the consonants in the suffix had been metathesized to *ks*. The 'x'-type invited paragogic *t*, providing a prototype for *betwixt*. Derivations from the 'x'-type include the 'xe'-type (e.g. *bitwexe*), 'xn'-type (e.g. *bitwixen*), 'xte'-type (e.g. *betwyxte*), 'xst'-type (e.g. *bitwixst*), and 'xts'-type (e.g. *betwixts*).

Thirdly, the 'betwēoh'-type ('h') is characterized by final *h*, which represents an old accusative ending of the numeral in a prehistoric period when the preposition governed the accusative. The *h*, however, was subject to loss in later times, and *-en* or *-es* tended to be appended on the analogy of variants that had such endings (e.g. *bitweien* and *bitweies*, an instance of the 's'-type). Other types derived from the 'h'-type include 'hn' (e.g. *bitwihan*) and 'he' (e.g. *bituhhe*).

Table 1 shows a diachronic distribution of the types of BETWEEN, as they can be distinguished in the *HC* data. It serves a starting point for the more detailed period-by-period description that follows.

Table 1. Distribution of historical variants of BETWEEN in the *HC*

Period/type	nm	nn	n	ne	x	xe	xn	xt	xte	xst	xts	h	hn	he	s
O2 (850–950)	14	1	–	–	13	–	–	–	–	–	–	31	–	–	–
O3 (950–1050)	5	22	15	–	56	–	–	–	–	–	–	49	–	–	–
O4 (1050–1150)	1	15	3	–	22	–	–	–	–	–	–	3	–	–	–
M1 (1150–1250)	–	28	3	8	14	–	1	–	–	–	–	–	5	8	–
M2 (1250–1350)	–	1	5	33	1	1	1	–	–	–	–	–	–	–	1
M3 (1350–1420)	–	–	5	44	29	42	2	1	–	4	–	–	–	–	–
M4 (1420–1500)	–	–	6	33	28	8	2	1	6	–	–	–	–	–	–
E1 (1500–1570)	–	–	12	76	2	–	–	28	4	–	–	–	–	–	–
E2 (1570–1640)	–	–	23	43	–	–	–	32	6	–	–	–	–	–	–
E3 (1640–1710)	–	–	54	8	–	–	–	14	–	–	1	–	–	–	–

Table 2 presents illustrative examples of each variant in contexts.

Table 2. Illustrative examples of each variant of BETWEEN

Type	Period	Text (line in *HC* file)	Context (keyword in italics)
'nm'	O2	COOROSIU(26301)	ðæt hie friþ him *betweonum* hæfden
'nn'	O3	COAELET3(34692)	And eac eow *betwynan* eowre fet aðweaþ mid eadmodnysse
'n'	O3	CODURHAM(23628)	ðætte god is fylgeþ gie *bituen* & in allum
'ne'	M1	CMANCRE(8435)	hit ne weoxe forþre *bitweone* mon & ancre
'x'	O4	COWULF4(15831)	Þes gedwolgod wæs arwurþe eac *betwux* eallum hæþenum on ðam dagum

Table 2. (*continued*)

Type	Period	Text (line in *HC* file)	Context (keyword in italics)
'xe'	M3	CMBOETH(55785)	For so as ther nis noon alliaunce *bytwixe* good folk and schrewes
'xn'	M3	CMWYCSER(16133)	to make diuision *bytwixen* hem and oðtre men
'xt'	M4	CMTOWNEL(7151)	I am agast that we get som fray *Betwixt* vs both
'xte'	M4	CMMALORY(17074)	for there was muche trw love *betwyxte* hem
'xst'	M3	CMHORSES(19203)	ðat it be fast sittynge *bi-twexst* ðe fyke & ðe hol skyn
'xts'	E3	CEPRIV3(61120)	there will come severall things *betwixts* that and the Abby
'h'	O2	COBOETH(67363)	ða yflan bioþ ungerade *betwuh* him selfum
'hn'	M1	CMHALI(31624)	hwuch schal beo ðe sompnunge *bituhen* ow ibedde?
'he'	M1	CMHALI(24596)	Hwet makeþ hit iluuet *bituhhe* beasteliche men
's'	M2	CMTHRUSH(702)	Hic herde a strif *bitweies* two

2.2 Period-by-period description[5]

I have not conducted a detailed investigation of OE in this study, but I refer to the diachronic distribution of the BETWEEN-types in the OE segment of Table 1. Kitson (1993: 11) collected from OE texts "a total of 1901 instances" representing "1010 spelling-variants" and localized these variants on an OE dialect map. One of his revealing findings was that *betweox* was "the exclusive form in north and west Wessex, and the most common in south and east England generally" (1993: 13–14). Kitson made no note of the 'xt'-type, nor did I find any instance of it when searching *The dictionary of Old English corpus* (*DOEC*); however, there are other references that testify to its presence.[6]

Moving on to the EME period, we can draw on *LAEME* to obtain the distribution of BETWEEN-types (see Table 3).[7]

---

5. Appendix 1 provides illustrative examples of BETWEEN (with source reference), as they occur in the corpora of different periods. Likewise, Appendixes 2 and 3 provide examples of AMONG and AMID, respectively.

6. Bosworth & Toller (1898) mention *betwyh, betwuht,* and *betwuxt* under the entry *betweoh*. Barnhart et al. (1973) also point out with regard to *betwixt* that "[t]he final t in *betwixt* developed in Old English." *DOEC* gives no instance of the 'xt'-type but dates *betuxt* as C13 and C14 under the dictionary entry "*be-twux* prep. and adv."

7. The 'hs'- and 'e'-types, unattested in the *HC* evidence, represent variants such as *twihs, be-tue,* and *bi-twe*.

Table 3. Distribution of historical variants of BETWEEN in the *LAEME*

Period/Type	nn	n	ne	x	xe	xn	xte	hn	he	s	hs	e
C12b[8]	18	1	2	7	–	–	–	–	–	–	–	–
C13a	23	4	19	7	5	4	–	9	14	–	–	–
C13b	20	5	23	2	1	4	4	1	–	1	2	1
C14a	5	16	28	10	2	2	–	–	–	–	–	1
Total	66	26	72	26	8	10	4	10	14	1	2	2

The distribution of BETWEEN-types in EME based on the *LAEME* squares well with the survey of the types in the *HC* (Table 1), but *LAEME* shows in greater detail the diachronic and diatopic distribution of the types in EME. Diachronically, the 'nm'-type had fallen out of use by the beginning of the EME period, as it was leveled to the 'nn'-type. The old 'n'-type remained present, but the 'ne'-type, with final schwa added, increased so much that it became the most common type throughout most of the period. In the meantime, the 'x'-type not only maintained itself, but it also developed the 'xe'-, 'xn'-, and 'xte'-types. Although the 'h'-type had fallen out of use by EME, the period saw a moderate development of the 'hn'- and 'he'-types, though not for a long time.

The diatopic distribution of variants also yields a number of insightful observations. In earlier work (Hotta 2014), I have provided a diatopic analysis of the *LAEME* evidence. Table 4 reproduces the distribution of variants of BETWEEN across the seven distinguished dialects, as presented in Hotta (2014: 27): N[orthern], N[orth] E[ast]M[idland], N[orth]E[ast]M[idland], N[orth]W[est]M[idland], S[outh]E[ast] M[idland], S[outh]W[est]M[idland], S[outh]W[estern], and S[outh]E[astern].

Table 4. Distribution of variants of BETWEEN across seven dialects of EME

Dialect	nn	n	ne	x	xe	xn	xte	hn	he	tn	tx	txn	txe	ths	s	e	Yn	zn	Total
N	–	–	1	9	2	2	–	–	–	–	1	–	–	–	–	–	–	–	15
NEM	14	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	14
NWM	7	–	6	–	–	–	–	8	14	–	–	–	–	2	–	–	–	–	37
SEM	14	20	9	5	–	–	–	–	–	3	–	1	–	–	–	–	–	–	52
SWM	31	1	26	7	5	7	–	2	–	–	1	–	1	–	1	–	1	1	84
SW	–	–	16	3	–	–	4	–	–	–	–	–	–	–	–	1	–	–	24
SE	–	–	14	–	–	–	–	–	–	–	–	–	–	–	–	1	–	–	15
Total	66	21	72	24	7	9	4	10	14	3	2	1	1	2	1	2	1	1	241

As Mustanoja (1960: 369) notes, the 'nn'-type occurred "mainly in the more southern parts of the country," but *LAEME* shows that it does appear to be common in

---

8. In these tables and below, century specifications are abbreviated as 'C12', 'C13', etc.; 'a' and 'b' stand for the first and second half of the century, respectively.

the North Midland dialects as well. Kitson's observation about the geographically restricted distribution of *betweox* in OE, as noted above, seems to square with the EME evidence in the *LAEME*, which shows that types including *x* occur as unique forms in the Northern dialects and as common forms in the South-West Midland dialects. The 'xte'-type, which will be seen to grow towards ModE, remained marginal at this stage, with variants like *bi-twixte* only occurring four times in the Southwestern dialects in the period C13b. As for the *h*-type, the *HC* evidence seems to indicate that it went out of fashion quickly after OE, but the *LAEME* evidence suggests that it survived well into EME despite its attestation being almost restricted to C13a North-West Midland.

In the LME period, the *MEGC* revealed the following distribution: 1 instance of the 'nne'-type, 62 instances of 'n', 216 of 'ne', 126 of 'x', 14 of 'xe', 2 of 'xn', 4 of 'xt', and 6 of 's'. The most common type remains the 'ne'-type. What characterizes this period most, however, is a remarkable growth in the relative share of the 'x'-type. As a result, other derivative types with *x* such as 'xe', 'xn', and 'xt' saw a slight growth as well. The 's'-type remained as marginal, as it was in EME.

To get an overview of the situation in EModE, I used *EEBO*. Since *EEBO* is not a compiled corpus but a collection of books published from 1473 to 1640, I compiled a text database of about 150 million words which I then used as a corpus. This custom-made database that I drew from *EEBO* was not as balanced and representative as more established corpora. It should also be noted that the size of the subcorpora for each half century varies greatly: 244,602 words for C15b; 3,277,691 for C16a; 13,166,673 for C16b; 48,784,537 for C17a; 83,777,910 for C17b; and 90,945 for C18a. In this and the following tables based on the *EEBO* evidence, figures are given in words per million (wpm) rather than in raw frequency. Most of the older distinct types had either fallen out of use or had been leveled by the EModE period, such that no more than three types were recognized: 'between', 'betwix', and 'betwixt'.

Table 5 clearly shows that in EModE the 'xt'-type far outnumbered the 'x'-type, although it hardly approached the ever predominant 'n'-type in frequency. The 'xt'-type seems to have reached its peak towards the end of EModE, perhaps in C17, and served as a respectable rival of the 'n'-type.

Table 5. Distribution of historical variants of BETWEEN in *EEBO* (figures given in wpm)

Period/Type	'between'	'betwix'	'betwixt'
C15b	196.24	36.79	69.50
C16a	239.68	1.825	36.20
C16b	175.59	4.025	57.87
C17a	245.51	0.57	123.13
C17b	256.12	0.084	85.17
C18a	197.92	0.00	164.93

The LModE distribution was obtained on the basis of the *CLMET3.0*, which contains about 34 million words, divided into three subcorpora of a roughly equal size at 70-year intervals (10,480,431 words for 1710–1780; 11,285,587 for 1780–1850; and 12,620,207 for 1850–1920). Figures are given in wpm. Only two types were attested in the corpus: 'between' and 'betwixt'.

Table 6. Distribution of historical variants of BETWEEN in the *CLMET3.0* (figures given in wpm)

Period/Type	'between'	'betwixt'
1710–1780	465.06	62.78
1780–1850	483.89	9.66
1850–1920	480.97	4.04

What Table 6 shows is that during the LModE period, the share of the 'xt'-type decreased dramatically, such that the 'n'-type was established as effectively the only form available. This distribution has continued to this day: a *BNCweb* search resulted in 920.52 wpm for *between* while it only yielded 0.31 wpm for 'betwixt'. In PDE, *betwixt* is nearly out of use.

2.3 Historical summary of BETWEEN

Throughout the history of English, variants of BETWEEN with *n* have remained the most common. The 'nm'-, 'ne'-, and 'n'-types were favored in OE, EME, and LME, respectively, and the last one of these, typically realized as *between*, has been a dominant type since EModE. The variants that contain *x*, including the 'xt(e)'-type, have been attested since OE, but among them the 'x'-type grew particularly in LME. In EModE, it was largely replaced by the 'xt'-type (realized as *betwixt*), which then reached its peak in C17, serving as the second most common variant. Nevertheless, *betwixt* became increasingly rare from LModE to PDE. All other types were either limited to particular periods (e.g. the 'h'-type largely to OE) or were used only sporadically (e.g. the 's'-type in ME).

3. AMONG

3.1 Etymological notes and historical overview based on the *HC*

PDE *among* has its origin in OE in the composite expression (where the meaning of each of the composite parts is clearly identifiable) *on gemang* "in the crowd", followed by the genitive of a noun. Towards LOE and EME, the reanalysis of the

second word as part of a lexicalized preposition along with phonetic reduction of the first component gave rise to variants such as *amang* and *imong*. In addition, forms with other prefixes like *bimong* emerged, as a result of analogy with *amang*. Variants with final *s* and final *st* first appeared in C13b and in C15b, respectively.

Variants of AMONG collected from the *HC* are classified into seven types: 'g', 'ge', 'ges' (the vowel varies between *e*, *i*, *y*, and *u*), 'gs', 'gest', 'geste', and 'gst'. Table 7 summarizes the results.

Table 7. Distribution of historical variants of AMONG in the *HC*

Period/Type	g	ge	ges	gs	gest	geste	gst
O2 (850–950)	2	–	–	–	–	–	–
O3 (950–1050)	14	–	–	–	–	–	–
O4 (1050–1150)	3	–	–	–	–	–	–
M1 (1150–1250)	37	2	–	–	–	–	–
M2 (1250–1350)	29	6	6	–	–	–	–
M3 (1350–1420)	80	5	23	1	–	–	–
M4 (1420–1500)	27	38	11	2	–	–	–
E1 (1500–1570)	19	43	3	2	8	2	15
E2 (1570–1640)	44	5	–	–	1	–	49
E3 (1640–1710)	57	–	–	–	–	–	37

The 'g'-type has been widely available throughout its history. The 'ge'-type with final *e*, first attested in EME, increased greatly in frequency towards C16 but afterwards went out of fashion. The 'ges'-type and, to a lesser extent, the 'gs'-type grew along with the 'ge'-type during ME, but in EModE they were displaced by the 'gest'-, 'geste'-, and 'gst'-types. The 'gst'-type, in its established form *amongst*, became so common during EModE that it even outnumbered *among* in C17. Table 8 presents illustrative examples of each variant in contexts.

Table 8. Illustrative examples of each variant of AMONG

Type	Period	Text (line in *HC* file)	Context (keyword in italics)
'g'	O3	COWSGOSP(58093)	ða ne for se H+alend na openlice *gemang* þam Iudeon
'ge'	M2	CMALISAU(19600)	*Amonge* hem of Perce was a kniȝth
'ges'	M3	CMBOETH(3422)	so that there ne be *amonges* hem no difference
'gs'	M4	CMROLLPS(4483)	i. eldyd *ymangs* all myn enmys
'gest'	E1	CEBIO1(28855)	It is therefore in lawe *amongest* Christen men insufficient
'geste'	E1	CEBOETH1(19604)	they shoulde be accountyd *emongeste* thynges that are to be desyred
'gst'	E1	CETRI1(88321)	the Holy Ghost be *amongst* you

3.2 Period-by-period description

Let us begin by looking at what the *LAEME* data tells us about the distribution of variants of AMONG in EME. The results are presented in Table 9.

Table 9. Distribution of historical variants of AMONG in the *LAEME*

Period/Type	g	ge	n	ges
C12b	22	–	–	–
C13a	68	1	2	–
C13b	95	3	3	4
C14a	49	5	–	3

The distribution based on the *LAEME* results is compatible with that based on the *HC*. Alongside the predominant 'g'-type (available throughout EME), the 'ge'-type (e.g. *imange*) as well as the 'ges'-type (e.g. *amanges*, *amangis*, and *amongus*, with different vowels before *s*) became available in C13b.

    With regard to LME, the *MEGC* survey shows that this period continued the trend observable in EME. Four distinct types occur with the following frequencies: 195 for 'g', 95 for 'ge', 38 for 'ges' (with different vowels), and 1 for 'gs'.

    By EModE, variants were reduced to three types: 'g', 'gs', and 'gst'. The distribution of the AMONG types in *EEBO* is presented in Table 10.

Table 10. Distribution of historical variants of AMONG in *EEBO* (figures given in wpm)

Period/Type	'among'	'amongs'	'amongst'
C15b	89.94	12.26	4.09
C16a	97.03	35.28	65.70
C16b	245.16	7.52	170.43
C17a	241.67	0.27	267.17
C17b	329.04	0.095	250.04
C18a	241.90	0.00	208.92

Most striking about the distribution of AMONG-types in EModE is that the 'gst'-type not only appeared for the first time in this period, but that it also saw such a dramatic growth during the period that by 1600, it had displaced the 'gs'-type from which it derived and even temporarily surpassed the ever steady 'g'-type in C17a. The 'gst'-type, however, declined, after it had reached its peak somewhere around 1650.

    Table 11, based on the *CLMET3.0* search, shows the following distribution in the subsequent LModE period.

Table 11. Distribution of historical variants of AMONG in the *CLMET3.0* (figures given in wpm)

Period/Type	'among'	'amongst'
1710–1780	387.39	65.46
1780–1850	396.52	136.28
1850–1920	438.19	78.68

The LModE period saw a continued decline of *amongst*, but less remarkable than that of *betwixt* (see above). Despite its steady fall in frequency over time, *amongst* has remained a viable alternative to *among*, whereas *betwixt* can hardly be regarded as a competitor to *between* today. The *BNCweb* search returns 227.64 wpm for *among* and 46.18 wpm for *amongst*.

3.3   Historical summary of AMONG

Among the historical variants of AMONG, the 'g'-type has been dominant through-out, except in C17 when the 'gst'-type outnumbered the 'g'-type, if only temporarily. The 'ge'- and 'ges'-types were first attested in EME texts and then increased in frequency towards C16; afterwards, however, they declined under the pressure of the innovative 'gst'-type. The EModE period was characterized by the striking growth of *amongst*, which competed with *among* in frequency and even surpassed it for some time around 1650. This peak, however, did not last long, and *amongst* soon took up second place again. Today *amongst* remains a viable, though more formal, alternative to *among*.

4.   AMID

4.1   Etymological notes and historical overview based on the *HC*

AMID can be traced back to OE *on middan* "in the middle". Like *among*, *amid* is in origin a composite expression made up of the preposition *on* and the dative noun *middan*, possibly followed by the genitive. This compositional makeup was already difficult to recognize in OE, in that *on middan* was treated as a lexicalized prepositional expression that governed the dative as well as the genitive. Despite its susceptibility to such reanalysis, the compositional character of the two elements lingered until LME, when separately written attestations such as *in mid* were still current.

I begin the historical survey of AMID by presenting the distribution of its historical variants, drawing on the *HC* (Table 12). Four types are distinguished: 'd' (e.g. *amid*, *inmid*), 'de' (e.g. *amidde*), 'ds' (e.g. *amids*), and 'des' (e.g. *amiddes*, *in-middes*, and *to middes*).

Table 12. Distribution of historical variants of AMID in the *HC*

Period/Type	d	de	ds	des
O1 (−850)	1	–	–	–
O2 (850–950)	6	2	–	–
O3 (950–1050)	12	1	–	8
O4 (1050–1150)	1	–	–	–
M1 (1150–1250)	3	1	–	–
M2 (1250–1350)	1	7	–	7
M3 (1350–1420)	1	4	–	8
M4 (1420–1500)	–	–	–	–
E1 (1500–1570)	–	–	1	–

The 'd'-type (e.g. *amid*) was widely present, particularly from OE to EME. LME is characterized by the predominance of the 'de'- and 'des'-types, both of which saw some sort of 'revival', since they were common in OE but unpopular in EME. The distribution of AMID-types in EModE suggests that all the types had become nearly extinct; indeed, the subperiods of E2 (1570–1640) and E3 (1640–1710) provide no attestation of any relevant variants. Further attestations can be found in the period-specific corpora below.

Table 13 presents illustrative examples of each variant in contexts.

Table 13. Illustrative examples of each variant of AMID

Type	Period	Text (line in *HC* file)	Context (keyword in italics)
'd'	O3	COBENRUL(66315)	Hæbbe se abbod *a mid* him gewrit ealra ðæra æhta
'de'	M1	CMKATHE(23523)	Her *amidde* wes ðis meiden iset
'ds'	E1	CEPLAY1B(7312)	how Hodg lieth tomblynge and tossing *amids* the floure
'des'	M2	CMEARLPS(23471)	Y ʒede in ðe innocens of myn hert, *amiddes* ðe wylle of myn hert

4.2   Period-by-period description

Let us first consider the EME distribution in *LAEME*, as presented in Table 14.

Table 14. Distribution of historical variants of AMID in the *LAEME*

Period/Type	d	de	den	t
C12b	–	–	1	–
C13a	5	4	3	1
C13b	2	13	2	–
C14a	1	14	–	–

The EME period saw a rise in frequency of the 'de'-type relative to its vowel-less rival 'd'; note that neither the 'ds'-type nor the 'des'-type was known yet. Despite the 'den'-type not being attested in the *HC*, *LAEME* records 6 instances of *amidden*, 4 from South-West Midland and 2 from North-West Midland. It would appear, then, that their distribution was limited dialectally and also diachronically because no other periods record any instance of it.

In the following period, LME, the 'des'-type, comprising *amiddes*, *emydes*, *in myddis*, was attested for the first time. Throughout LME, this was a common type, alongside the 'de'-type. The *MEGC* search did not reveal any variant with *st*, but Wełna (2014: 333) reports the unique ME attestation of *emyddiste* in the *Alphabet of tales* c. 1450. The results of my *MEGC* search are as follows: no instance of 'd', 9 of 'de', and 9 of 'des'.

The following period is EModE; the distribution of AMID-types based on *EEBO* is shown in Table 15.

Table 15. Distribution of historical variants of AMID in *EEBO* (figures given in wpm)

Period/Type	'amid'	'amids'	'amidst'
C15b	0.00	8.18	0.00
C16a	0.00	1.22	0.00
C16b	9.87	1.75	9.34
C17a	1.46	0.67	6.76
C17b	0.19	0.048	9.80
C18a	0.00	0.00	0.00

As we saw earlier, little or no evidence was found in the *HC* of AMID-types in EModE, but *EEBO* contains relevant data. Table 15, however, shows that we are still faced with a relative data shortage for this period. What little data there is shows indeed that the 'amids'-type first appeared towards EModE, but it is apparent from the figures that AMID remained relatively infrequent in any variant in the subperiods C15b and C16a. At the same time, corpus sizes for the earliest subperiods are relatively small, so the limited attestation of the lexical item in C15b and C16a may be more apparent than real. Such being the case, it is striking, when we look at

the evidence from 1551 onwards, that the 'amid'-type re-emerged long after it had been relatively dormant in ME. Even more striking is the fact that the innovative 'amidst'-type with paragogic *t* not only appeared rather suddenly in the second half of C16 but also easily and rapidly surpassed the 'amid'-type in C17. The lack of evidence from C18a is again supposed to be due to the small size of the subcorpus.

With regard to the LModE period, the *CLMET3.0* confirms our earlier observation that by 1700 the 'amidst'-type had become more common than the 'amid'-type. As Table 16 shows, the predominance of *amidst* over *amid* continued throughout C18, but apparently the relative share of *amid* surpassed that of *amidst* again in the course of the 19th century.

Table 16. Distribution of historical variants of AMID in the *CLMET3.0* (figures given in wpm)

Period/Type	'amid'	'amidst'
1710–1780	2.67	38.17
1780–1850	33.31	40.94
1850–1920	50.71	14.10

Today *amid* is far more common than *amidst*, with 10.83 wpm for the former and 4.92 wpm for the latter in the *BNCweb*. This gap can be considered to represent an extension of the late C19 distribution. It is to be noticed, however, that the lexical item is considerably rarer today than it was in LModE.

4.3   Historical summary of AMID

The AMID-types enjoyed varying prevalence across the various periods discussed: the 'd'-type was predominant from OE to EME, the 'de'-type in EME, and the 'des'-type in LME. In the transition period from LME to EModE, we are faced with data scarcity, most likely due to the small corpus size; still, available evidence suggests the emergence of the 'amids'-type as well as the apparent disappearance of the older types. After this period, from 1551 onwards, more evidence became available showing that the innovative 'amidst'-type became highly frequent, exceeding the old 'amid'-type in frequency in C17. After its peak in C18, *amidst* gradually declined towards 1900 as a result of competition with a revitalized *amid*. Since then the distribution of the two alternatives has been relatively stable, although AMID has become rarer than before.

Figure 1 summarizes the complex ebb-and-flow distribution of the variants of BETWEEN, AMONG, and AMID. The asterisks and dotted lines (each representing a decade) denote that the variant is relatively common and uncommon, respectively. The periodization is to be interpreted as approximate.

		C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21
	`nm`		*****	*****-----	----------	-----									
	`nn`		-----	-----*****	**********	**********	*******---	-----							
	`n`			-----	----------	----------	----------	----------	----------	-------***	**********	**********	**********	**********	*
	`ne`					-----	--*******	**********	**********	**********	****------	-			
	`x`		*****	**********	**********	**********	****-----	-----****	**********	----------					
	`xe`						-------	-----****	**********	-------				
	`xn`				-----	----------	----------	----------							
BETWEEN	`xt`						-----	----------	----------	**********	*******---	- - -	- - -	- - -	
	`xte`					---		-------	----------	----					
	`xst`						-----	--							
	`xts`										-------				
	`h`		*****	**********	*****-----	-----	-----								
	`hn`					-----	-----								
	`he`				-----	-----									
	`s`						-----	----------	-----						
	`g`		*****	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	**********	*
	`ge`					-----	----------	----------	--*******	*******---	----				
	`ges`						-----	-----*****	**--------	-------					
AMONG	`gs`						-----	----------							
	`gest`									----------	----				
	`geste`									-------					
	`gst`								-----	-------***	**********	****------	----------	----------	-
	`d`	----------	-----*****	**********	*****-----	----------	----------	----------	--	*****	----------	-----*****	**********	**********	*
	`de`		-----	----------	----------	----------	-----*****	**********	**						
AMID	`des`			*****	*****		*****	**********	**						
	`ds`								-----	----------	----------				
	`dst`								-	*****	**********	**********	**********	**--------	-

Figure 1. Diachronic summary of variants of BETWEEN, AMONG, and AMID

5.   Discussion

5.1   Proposed accounts of final *t*

One of the major questions that the present study raises is how and why *t* was appended to the end of the words, producing such variants as *betwixt*, *amongst*, and *amidst*. Several views have been proposed about this process, and they are summed up well in the *OED* and Dobson's comments. The following is what *OED* has to say in the entry for *against*.

> The development of excrescent final -*t* … was probably reinforced by the fact that the word was frequently followed by *te*, variant of THE *adj.*, and perhaps also by association with superlatives in -*st*; compare similarly AMONGST *prep.* 1a, AMIDST *adv.*, BETWIXT *prep.*   (*OED*, s.v. *against*)

Dobson's (1968: Section 437) comment is much more detailed, but it makes the same points as the *OED*. The various views presented by the *OED* and Dobson may be subsumed under three headings: (i) the addition of *t* is a phonetically motivated process, or paragoge; (ii) it concerns word boundary morphonology, in which the dental that originally belonged to the definite article typically following the preposition is attracted to the end of the preposition so that it may be reanalyzed as part of the preposition; and (iii) *t* was appended to these prepositions because they constituted a small lexical group of function words with a semantic component associated with the superlative. In the following sections, I examine each account and then introduce additional points of view to sketch a likely scenario of the development of the function words with final /st/.

5.2   Paragoge

Before it is possible to discuss paragogic *t*, I first consider the addition of *s* because it is a prerequisite for the paragoge of *t*. As was mentioned in 2.1, final *s* in several function words can be traced back to the genitive ending of their component nouns, which attributes an adverbial (and then prepositional or conjunctional) function to those nouns. Several words have been created in this fashion: *besides*, *needs*, *nowadays*, and *sometimes*.

    To this genitive-derived *s* was added paragogic *t*, yielding *betwixt*, *amongst*, and *amidst*. Many scholars believe the insertion of *t* to be a phonetic whim that applies to some words but not to others; others maintain that the addition of *t* can be attributed to morphological or lexical factors. In that respect, Wełna (2014: 336), for instance, writes: "In most cases the presence of the voiceless dental stop can be

due to the influence of morphological factors, like in the case of the adoption by words ending in *n* of the common suffixes from words of Latin or French origin, e.g. *ancient*, etc., into the vocabulary of English." Such morphological or lexical factors may be thought to include the shared final /s/ in function words like *betwix*, *amongs*, and *amids*. In my view, however, the addition of /t/ started as a haphazard phonetic process, firstly because it applied to words that apparently had no shared formal or functional characteristics (e.g. *against*, *behest*, *ernest*) and secondly because variants of the prepositions seem to have begun inserting *t* at different times: *betwixt* in C13 (or possibly in OE) and *amongst* and *amidst* both in C15. Perhaps they developed paragogic *t* independently when they did; this does not mean, however, that the resultant final /st/ may not have helped to develop a sense of lexical association between the prepositions at a later stage. I return to this matter shortly.

5.3   Word boundary morphonology

A second justification of the addition of *t* is in terms of word boundary morphonology and involves a kind of euphony at the word boundary and subsequent reanalysis. It is phonetically plausible that [θ] or [ð], the initial sound of the definite article or demonstratives, turned to [t] immediately after the word-final [s] of the function words (e.g. *amids*) and that the resultant [t] was then reanalyzed as the ending of the preceding word.

How can this account be supported by historical evidence? One way would be to single out the period when the types with final /st/ began to be used commonly at the expense of the types with final /s/ and, for that period, to count the number of tokens of the /st/-types that were followed by dentals such as [θ] and [ð]. In the case of BETWEEN, the 'betwixt'-type increased dramatically in EModE, so this period could be a testing ground for the word boundary effect. In Hotta (2014: 30–31), concordance lines obtained from the EModE period of the *HC* were tested for the word boundary effect. The results showed that [θ] or [ð] (or [t] because of its being equally subject to reanalysis) following BETWEEN were in fact very common (roughly half of the relevant instances), which seemed to confirm our expectations.

Although the word boundary effect seems to be a promising explanation, there are two problems with it. Firstly, it is paradoxical that the supposed euphony around the word boundary facilitated types with /st/ because the resultant sequence /st/ followed by a consonant (very often plus the [θ, ð, t] of definite articles) would be extremely difficult to articulate, assuming all the consonants would be properly pronounced. It is as if euphony today led to cacophony tomorrow. Secondly, the word boundary account concerns the process, rather than the result, of euphony. It is uncertain how much of the dynamic process would have been reflected in writing.

In fact, there are no attestations in writing such as *betwixt e* for *betwixt the*, which would signal that the dental was extracted from the definite article and assigned to the preceding function word.

Despite the arguments against the word boundary account, I suspect that the proposed morphonological process is in fact plausible. For the moment, however, for lack of additional evidence, it seems safest to reserve it as supplementary to the account of haphazard and independent paragoge, as argued for in the foregoing section.

5.4   Semantic association with the superlative

Unlike the previous two accounts, the third is functional, involving a semantic as well as a morphological association of final /st/ with the superlative. As we saw earlier, the *OED* speaks of the "association with superlatives" as a possible account for the paragogic [t], which results in a [Cst]-ending ('C' standing for any consonant) in words such as *against*, *amidst*, *amongst*, *betwixt*, and *whilst*. On the other hand, Dobson (1968: Section 437) argues against this view, saying "there seems no good reason why the superlative should exercise the influence alleged." Despite Dobson's observation, there is evidence in favor of the association with the superlative.

The fundamental function of the superlative is intensification in degree; it is therefore not unlikely that final /st/, clearly associated in form with the superlative, should be reinterpreted as a marker of intensification of a more general nature, such that *betwixt*, *amongst*, and *amidst* can be taken to mean "right in between", "in the very middle of", and so on. Admittedly, these function words may not contain the notion 'degree' in their semantics, as gradable adjectives or adverbs do. However, the sense of 'middleness' that they share may convey a sense of contrast, if not degree, as the intensification of 'middleness' arises in everyday usage in contrast with spatial front and back or with temporal or sequential first and last. Reinforced phonetically by the coda /st/, the item *midst*, for instance, may be associated with the superlative sense approaching "middle in its most precise sense". To make my point clearer, the following quotations in the *OED* and *EEBO* are helpful: "In the whiche gospel it is profitable to men desyrynge God, so to knowe the first, the mydmeste, other the last" (*OED*, s.v. *midst*, Wycliffe's *Bible*, Early Version, c. 1384) and "On Earth joyn all yee Creatures to extoll Him first, him last, him midst, and without end" (*OED*, s.v. *midst*, Milton, *Paradise Lost*, 1667). A similar example is the following: "we may see that the poet hath studied to place the most generall in the first place, and the next generall in the midest, and the most speciall last of all" (*EEBO*, *The logike of the moste excellent philosopher P. Ramus Martyr, Newly translated, and in diuers places corrected, after the mynde of the Author*, 1574).

Further support for the transfer between the intensifying/superlative sense of *midst* and *mydmeste* onto the function word AMID comes from the many instances of AMID attested in LME and EModE such as *in þe myddis*, *in myddes of*, *in ye midst of*, and *in the very midst* for AMID. It is to be remembered that these were the periods when hardly any tokens of AMID could be attested in the data (see Section 4.2). The few instances of AMID in LME and EModE certainly indicate a continued presence of non-phrasal (or simplex) tokens of AMID, but evidence seems to show that in these periods there was a preference for prepositional phrases equivalent to AMID rather than the simplex AMID. Under these circumstances, it is not unlikely that the noun *mid(s)(t)* (with final /st/ assumedly deriving from the superlative of the adjective *mid*) was felt closely associated both formally and functionally with the preposition *amid(s)(t)*. This association may not necessarily have triggered the use of the prepositional variants ending with *st*, but is likely to have encouraged the use of these forms once they became available as a result of phonological processes.

5.5 Small lexical group of function words

Let us now try to integrate the different views into a historical scenario. The formal and functional processes described in the foregoing sections started independently but later merged with one another. The times when variants with final /st/ were first attested in each of the lexical items vary from OE to LME, but their adoption as common variants coincided in the 16th century. Another related concurrence is the common decline of the rival types with /s/ somewhere between the 16th and 17th century. These coinciding processes suggest that there was such a strong association, both formally and functionally, between the lexical items in question that they constituted a small lexical group of function words. The features tying them together were formally final /st/ and functionally the sense of 'middleness', as well as their being function words. All the features, taken together, contributed to developing a mutual lexical association.

Once the lexical connection was established, the morpheme-like unit *st* developed in this small lexical group, which then contributed to establishing or creating a number of other *st* words such as *against*, *whilst*, and *unbeknownst*.

5.6 Accounting for the decline of the types with /st/

So far, our discussion has been mostly concerned with the emergence, development, and establishment of the types with final /st/. In this section, I take up their decline on the path to PDE after they had enjoyed great popularity in ModE. The chronology and extent of the decline differed from one lexical item to another:

betwixt started to decline in C17b and is very rare now (0.31 wpm against 920.52 wpm for *between* in *BNCweb*); *amongst* in C18b, and it is now still frequent (45.18 wpm against 227.64 wpm for *among* in *BNCweb*); and *amidst* in C20, being now fairly common (4.92 wpm against 10.83 wpm for *amid* in *BNCweb*).

The difference in frequency between the types with and without /st/ might be related to their semantic differentiation. Burchfield (1998: 48) refers to the *OED*, which points out "a tendency to use *amidst* more distributively than *amid*, e.g. of things scattered about, or a thing moving, in the midst of others," but admits that "the distributional pattern of the words is not clearly ascertainable." Alternatively, the difference in frequency, say, between *amongst* and *among* might be attributed to their phonetic sensitivity to the presence or absence of a following vowel. Burchfield, however, notes that "[a]n older view, favoured by Fowler, that *amongst* is more common than *among* before a following vowel does not seem to be borne out by the evidence." It appears, then, that neither a semantic nor a phonotactic explanation seems to account for the difference in frequency of the types with and without /st/, much less the decline of the types with /st/.

Stylistic considerations, on the other hand, provide a more promising explanation for the decline of the types with /st/, since in PDE they clearly have a more formal, literary, or archaic ring than their counterparts without /st/. What we need to understand is the way the types with /st/ acquired such stylistically marked features.

In the present study, I have not conducted a stylistically oriented analysis on the attestations drawn from the historical corpora, and therefore I am not in a position to decide whether the acquisition of stylistic markedness coincided with the decline in frequency. Here, I tentatively propose a phonotactic factor that may have helped to produce an elevated style. Since EModE, the word-final three-consonant sequence [Cst] has become increasingly rare as a result of the loss of *-(e)st* as an inflectional suffix for the second-person singular verb along with the loss of the pronoun *thou* itself. In fact, in Standard PDE, there are very few commonly used words that have final [Cst] (e.g. *angst*, *next*, and *text*), except the preterit and the past participle of the verb that ends with [Cs] such as *fixed* and *glimpsed* and the words under study (e.g. *against*, *amidst*, *amongst*, *betwixt*, and *whilst*). Today final [Cst] sounds so rare, so distinct, and perhaps so strongly associated with the conservative *thou didst*, *thou canst*, and the like that it may be exploited for stylistic effects. I leave this possibility open, however, until experimental work is conducted along these lines.

6.   Conclusion

I suggest the following scenario for the ebb and flow of the types with final /st/ of BETWEEN, AMONG, and AMID. From OE to ME, there was a wide range of variants coming and going, some becoming more frequent than others. The sources of most variants were morphonological processes such as the insertion of genitive *-(e)s*, the leveling of word final segments, and analogical formation on existing variants. One important process, from the present point of view, was the insertion of final /t/ after /s/, producing new variants like *betwixt*, *amongst*, and *amidst*. These paragogic developments were, however, phonetically haphazard and independent in nature, and therefore variants with /st/ remained minor alternatives for a while.[9] The word boundary effect might be taken as a facilitator of the paragoge, but it is difficult to evaluate on the basis of historical evidence.

Towards LME, a form–function association between final /st/ and the superlative emerged, partly due to the shared sense of 'middleness' of the stems and partly through the close connection between the noun *midst* and the preposition *amidst*. As a result, the three lexical items could be considered to make up a small group of function words, which also began to behave similarly. For all three, the /st/-types became dominant almost simultaneously in C16 at the expense of the types with /s/, which declined subsequently from C16 to C17. Although the /st/-types were highly frequent for some time during the ModE period, they gradually became stylistically marked towards PDE, perhaps partly because their three-consonant cluster had a formal, literary, or archaic connotation.

This scenario needs to be supported by further evidence, but it fits at least with evidence from historical corpora. Further study of these issues may involve the following: (i) examining how stylistic markedness was achieved from ModE to PDE;

---

9.   If we look beyond modern standard varieties of English, there are abundant examples of *t* being inserted or lost after *s*. In his *EDD*, Wright (1968 [1905]: Section 295) gives examples from various dialects in the British Isles, including *beas(t)*, *fas(t)*, *hoarse(t)*, *ice(t)*, *jois(t)*, *las(t)*, *nex(t)* (after [s]), *nice(t)*, *once(t)*, and *twice(t)* (in all these examples, *t* is inserted after [s]). One of the reviewers comments that describing the loss or insertion of *t* after *s* as the result of a "haphazard phonetic" process does not really solve the problem much. To this comment, I would like to reply that initial phonetic motives behind the addition of *t* to the prepositions in question were likely the same as for other items such as *behest* and *ernest*, i.e. the phonotactic context where the final segment is *s*. Such an addition must have been 'haphazard' in the sense that the resultant form with final *t* might or might not be established as a variant later on, if it was phonetically motivated when first brought into being. I stress that, besides forms with final *t*, there were a host of variants with different endings. The question of why such variants emerged in the first place is almost always difficult to answer, though curious, and one tentative attempt would be a 'haphazard phonetic' process.

(ii) studying other lexical items that can be considered to belong to the same lexical group, such as *against*, *whilst*, and *unbeknownst*; and (iii) exploring the distribution of the prepositions in question in OE.

Acknowledgements

References

Corpora and dictionaries

ASD = Bosworth, Joseph & T. Northcote Toller (eds.). 1898–1921. *An Anglo-Saxon Dictionary*. Oxford: Clarendon.

Barnhart, Clarence L., Sol Steinmetz & Robert K. Barnhart (eds.). 1973. *The Barnhart dictionary of New English since 1963*. Bronxville, NY: Barnhart.

BNCweb = Hoffmann, Sebastian & Stefan Evert (eds.). 1996–2008. *BNCweb CQP-edition*. http://bncweb.lancs.ac.uk/ (2 February, 2015.)

CLMET3.0 = De Smet, Hendrik, Hans-Jürgen Diller & Jukka Tyrkkö (eds.). 2013. *The corpus of Late Modern English texts, version 3.0*. Leuven: Department of Linguistics, University of Leuven.

DOEC = Healey, Antonette diPaolo, Joan Holland, Ian McDougall & Peter Mielke (eds.). 2000. *The dictionary of Old English corpus in electronic form, TEI-P3 conformant version, 2000 release*. CD-ROM. Toronto: DOE Project 2000.

EDD = Wright, Joseph (ed.). 1898–1905. *The English dialect dictionary*, 6 vols. Oxford: Henry Frowde.

EEBO = Early English books online. ProQuest LLC. 2003–. http://eebo.chadwyck.com/home (2 February, 2015.)

eLALME = Benskin, Michael, Margaret Laing, Vasilis Karaiskos & Keith Williamson. 2013. *An electronic version of A linguistic atlas of late mediaeval English*. Edinburgh: University of Edinburgh. http://www.lel.ed.ac.uk/ihd/elalme/elalme.html (2 February, 2015.)

HC = Rissanen, Matti, Merja Kytö, Leena Kahlas-Tarkka, Matti Kilpiö; Saara Nevanlinna, Irma Taavitsainen; Terttu Nevalainen & Helena Raumolin-Brunberg (eds.). 1991. *The Helsinki corpus of English texts*. Helsinki: Department of Modern Languages, University of Helsinki.

LAEME = Laing, Margaret & Roger Lass (eds.). 2007. *A linguistic atlas of early Middle English, 1150–1325*. Edinburgh: University of Edinburgh. http://www.lel.ed.ac.uk/ihd/laeme1/laeme1.html (2 February, 2015.)

LALME = McIntosh, Angus, Michael Louis Samuels & Michael Benskin 1986. *A linguistic atlas of late mediaeval English*. With Margaret Laing & Keith Williamson. Aberdeen: Aberdeen University Press.

MED = Kurath, Hans, Sherman McAllister Kuhn & John Reidy (eds.). 1962–2001. *Middle English dictionary*. Ann Arbor: University of Michigan Press.

MEGC = Stenroos, Merja, Martti Mäkinen, Simon Horobin & Jeremy Smith. 2011. *The Middle English grammar corpus, version 2011.1*. Stavanger: University of Stavanger. http://www.uis. no/research/culture/the_middle_english_grammar_project/ (2 February, 2015.)

OED = Simpson, John & Edmund Weiner (eds.). 1989. *Oxford English Dictionary*, 2nd ed. Oxford: Oxford University Press.

Secondary sources

Burchfield, Robert (ed.). 1998. *Fowler's modern English usage*. Revised 3rd edn. Oxford: Oxord University Press.

Ciszek-Kiliszewska, Ewa. 2013. Middle English preposition and adverb *twix*. Token 2. 75–91.

Dobson, Eric John. 1968. *English pronunciation 1500–1700*, vol. 2, 2nd edn. Oxford: Oxford University Press.

Hotta, Ryuichi. 2014. Betwixt and between: The ebb and flow of their historical variants. *Journal of the Faculty of Letters: Language, Literature and Culture* 114. 17–36.

Kitson, Peter. 1993. Geographical variation in Old English prepositions and the location of Ælfric's and other literary dialects. *English Studies* 74. 1–50. doi:10.1080/00138389308598839

Mustanoja, Tauno Frans. 1960. *A Middle English syntax*. Helsinki: Société Néophilologique.

Wełna, Jerzy. 2014. Insertion and loss of the voiceless dental plosive [t] in Middle English. In Michael Bilynsky (ed.), *Studies in Middle English: Words, forms, senses and texts*, 329–342. Frankfurt am Main: Peter Lang.

Wright, Joseph. 1968 [1905]. *The English dialect grammar*. Oxford: Oxford University Press.

Appendix 1

Examples of BETWEEN (with reference), as they occur in the corpora of different periods.

LAEME: "dA SPAK+EN HIE HEm *BETwIENEn*" ('nn'-type; in the text vvat (4351) from C13a (SEM)); "*BI-TWEN* HIS HOND+ES HE BAR+ +IT" ('n'-type; in the text havelokt (7087) from C14a (SEM)); "HE BE+GIN+y TO GROCH+I *BE-TUENE* HIS TEy" ("ne"-type; in the text ay-enbitet (30324) from C14a (SE)); "*yER-EFTer waeX SUYTHE MICEL UUERRE *BETUYX* yE KING & EORL" ('x'-type; in the text petchront (1853) from C12b (SE)); "*IC HAF A+ +DERNE Pr^IUETE *TO SCHEW *BYTUIXE* ME AnD yE" ('xe'-type; in the text edincmbt (26504) from C14a (N)); "yA RICH+AN yE RIHT+LICHE LIBB+Ad MAgEN BEO+N *BITwIXEN* GOD+ES wRECCH+An" ('xn'-type; in the text lamhomA1t (18449) from C13a (SWM)); "*FOR MAK+IE LOFE *BI-TWIXTE* HEOM" ('xte'-type; in the text layamonBOt (10606) from C13b (SW)); "BITUHEN HARD & NESCHE *BITUHEN* wA OF yIS wORLD & TO MUCHEL wUNNE *BITHUEN* MUCHEL & LUTEL" ('hn'-type; in the text royalkgat (2301) from C13a (NWM)); "yurh A+ +SCHENE SCHADEwE *BITUHHE* ME & HIM" ('he'-type; in the text royalkgat

(4013) from C13a (NWM)); "*HIC HER+DE A STRIF *BITWEIES* TWO" ('s'-type; in the text digby86mapt (12857) from C13b (SWM)); "LOF AND PES *TWIHS* GODD AND MAn" ('hs'-type; in the text bodley26t (442) from C13b (NWM)); "PAYS BE-TUENE GOD AND MAN BE-TUENE MAN AND ANGLE *BE-TUE* MAN AND HIM+ZELUE" ('e'-type; in the text ayenbitet (39018) from C14a (SE)).

MEGC: "HE WAS AN-HONGED ON yE RODE * *BYTWENENE* TWO yEOUES" ('nne'-type; in the text Gloucs_L6980_OK1 (238) from Gloucestershire); "THE WERRE ENDURED *BE-TWEN* HEM IJ zERE" ('n'-type; in the text Essex_L9250_OK1ms (304) from Essex); "GRETE DEBATE WAS THAT SAME TYME *BETWENE* THE EMPerOur & POPE" ('ne'-type; in the text Herefs_L7481_OK1 (25) from Herefordshire); "yE TREW ENTENT OF yE CONCLUSION@ *BETWYX* yE SAYDE ParTYES" ('x'-type; in the text Lancs_L0420_OK1 (25) from Lancashire); "*GRET DIFFERENCE IS yERE *BE-TWIXE* PEYNES" ('xe'-type; in the text Norfolk_L4648_OK1 (281) from Norfolk); "GOSTLY DETH I%S A DEParTENGE THOUR SENNE *BY-TWYX-EN* MANES SOWLE GRETE GOD : & MANES SOULE W\T-ENNE" ('xn'-type; in the text Suffolk_L4266_OK2 (251) from Suffolk); "*BYTWYXT* THEM̃SHAL+BE GRETE MURTHER" ('xt'-type; in the text Somerset_L5171a2_OK2ms (52) from Somerset); "HIT YS A-CORDYT AGRET & APOYNTYT *BETWYS* YE SAID ELIZABETH AND *JOHN YE SONn" ('s'-type; in the text Cumb_L0535_OK1 (8) from Cumberland).

EEBO: "there is commonly as much difference *between* Trade and lending of mony, as *betwixt* a Trades man borrowing of mony, and one that is no trades man lending it." ('n'- and 'xt'-types; in the text *The PRETENDED PERSPECTIVE-GLASS …* (1664)); "I Put Crystis passyon *betwix* me & myn euyll werkes / and *betwix* me & his wrathe" ('x'-type; in the text *Here Begynneth a lytell treatyse called Ars moriendi* (C15b)).

CLMET3.0: "I was forced to set out *between* Five and Six o'Clock in the Morning" ('n'-type; in the text *A narrative of the life of Mrs. Charlotte Charke* (1755)); "Is there any comparison *betwixt* you and Caelia?" ('xt'-type; in the text *The school for lovers* (1762)).

Appendix 2

Examples of AMONG (with reference), as they occur in the corpora of different periods.

LAEME: "BLISC+ED BE yU *MANG* ALL+E WIMMEM" ('g'-type; in the text cotcleoBvit (248) from C13a (N)); "*CRIST+ES HELPE BE US *A-MONGE*" ('ge'-type; in the text genexodt (5784) from C14a (SEM)); "yER-HEFTER COM URE LOUERD *AMOn+* +HAM ALL+E STOnD+EN" ('n'-type; in the text tr323at (6116) from C13b (SWM)); "*GOD+IS WORD TO SPEL AND SPRED+E *AMAnGIS* ALL+E MIS+LEU+AnD LEDE" ('ges'-type; in the text edincmct (10135) from C14a (N)).

MEGC: "THE CHYLD SAYDE SYNNYS FYUE * *A-MONG* MAN-KYNDE" ('g'-type; in the text Suffolk_L8491_OK1ms_base from Suffolk); "ALL THYNG WAS DONE BOTHE *AMON-GE* HYE AND LOWE" ('ge'-type; in the text Herefs_L7481_OK1_base from Herefordshire); "DEBATE SCHULD EUerMORE BIEN *A-MONGES* VS & HEM" ('ges'-type; in the text Essex_L9360_OK1ms_base from Essex); "YerFORE THIS *CLOTON HADE NO MORE LOND *AMONGS* HOM" ('gs'-type; in the text Staffs_L0227_OK1_base from Staffordshire).

EEBO: "It was a lawe *among* the Lacedemonians" ('g'-type; in the text *A SERMON PREACHED AT PAVLES CROSSE THE IX. OF NOVEMBER, 1589* (1589)); "it was ordeined and enacted *amonges* other thinges" ('gs'-type; in the text *A proclamation deuised and made by the kynges highnes … (1544)*); "which hath caused a great feare *amongst* the Souldiers on all sides" ('gst'-type; in the text *THE CERTAINE AND TRVE NEWES … (1621)*).

CLMET3.0: "Beattie is *among* the philosophers what the Quaker is among religious sectaries" ('g'-type; in the text *Lives of English Poets* (1846)); "There had, in fact, always been *amongst* them a conspiracy against me" ('gst'-type; in the text *Memoirs of Henry Hunt* (1820–22)).

Appendix 3

Examples of AMID (with reference), as they occur in the corpora of different periods.

LAEME: "MANI MIRTH+ES ER *E-MEDD* YE LEUE+LE+IST OF ALL LAND+ES" ('d'-type; in the text cotvespcmat (7728) from C14 (N)); "HE COM *AMIDDE* yE PUTTE *yE WOLF" ('de'-type; in the text digby86mapt (15936) from C13b (SWM)); "MAHTE HABB+EN HARE BREAd wId AL HIS OdER wA RIHT *AMIDDEN* HIS NEASE" ('den'-type; in the text corpart (16600) from C13b (NWM)); "*ME wEORP HAM MIT TET-ILKE *AMIT* TE LEIE yER" ('t'-type; in the text royalkgat (10405) from C13a (NWM)).

MEGC: "*SOME CLERKES SEYN AS BOKES BERIY WITNES yAT HELL EUENE *AMYDDE* yE EORyE YS" ('de'-type; in the text Wilts_L5420_OK1_base (333) from Wiltshire); "yE TRE OF LIFE yAT IS *IN MYDDIS* PARADISE" ('des'-type; in the text Lincs_L0422_OK1_base (336) from Lincolnshire).

EEBO: "Meete with the Queene *amid* the way she came" ('d'-type; in the text *ELIZABETHA TRIVMPHANS … (1588)*); "And when I am at ease, *amids* my pleasant haps" ('ds'-type; in the text *A discourse of the great crueltie of a Widowe … (C16b)*); "The wonted roare was up *amidst* the woods" ('dst'-type; in the text *A MASKE PRESENTED At Ludlow Castle, 1634* (1637)).

CLMET3.0: "She accused him, *amid* sobs, of putting his cousin before his wife and son" ('d'-type; in the text *The Old Wives' Tale* (1908)); "He seemed to stand firm *amidst* the confusion and terror" ('dst'-type; in the text *Olive* (1850)).

CHAPTER 10

English word clipping
in a diachronic perspective

Donka Minkova

UCLA (University of California Los Angeles)

Shortening processes are increasingly productive in Present-day English (PDE), generating 9–15% of the new words in the language. They are therefore a valuable source of information on the interaction of production, perception, prosodic well-formedness, morphological patterning, and sociolinguistic and pragmatic aspects of word-formation. One type of shortening, word clipping, is considered unpredictable by some scholars (Durkin 2009: 116; Don 2014: 27), while others (Plag 2003; Lappe 2007; Berg 2011) find non-random patterns in the outcome of clippings in PDE. The diachronic records of clipping have not been addressed systematically in the literature. Data collected from the online *OED* show that clipping in early English is restricted to fore-clipping, which peaked between 1300–1600 and then decreased quite sharply. Back-clipping was practically unattested until the end of Middle English, whereupon it rapidly became the dominant model, making fore-clipping a marginal pattern for common words. The association between prefixation and lack of stress during the period of widespread fore-clipping, and the stability of the left edge of the word in back-clippings after the 15th century require reference to both the surplus location and composition, and the residue structure. An account in terms of ONSET, STRESS, Morphological identity, and ALIGN-L shows how the relative importance of these constraints changes in tandem with independently established changes in the history of English.

Keywords: clipping, stress, onset, morphological identity, alignment, aphesis

1. Definition of terms – Scope of coverage

This study describes the diachronic patterns of 'clipping' in Old and Middle English and compares them to the patterns observed after the 16th century. Clipping is a subtype of 'shortening', as is 'truncation' (and many other processes such as abbreviations and acronyms, blends, and back-formation). The term 'truncation' is avoided here as there are sources in which it is used as a cover term for *any* shortened form, or

where it refers to hypocoristic forms (names). In the usage adopted here, clipping is reserved specifically for "truncated words which are not personal names," as in Lappe (2007: 4). The clarification is needed because the application of the terms varies depending on the analysis. Bauer et al. (2013: 404) consider the two types "distinct morphological categories." Berg's (2011: 16) position is that "[i]n much the same way as foreclippings and backclippings are different instantiations of the same phenomenon called clipping, the shortening of common nouns and first names should be understood as a unitary process." He attributes the partially different strategies of shortening for common nouns and names not to morphological but to pragmatic factors, "disparate levels of availability," while Jamet (2009: 16) considers the two terms synonymous. Bauer & Huddleston (2002: 1635) do not subsume (nick)names under clipping because they "often introduce extraneous complications." The term 'clipping' itself was first recorded in 1933 – the *OED* attributes the coinage to Eric Partridge. Jespersen's own term *stump-words* (1922: 169–171) did not catch on.[1]

Another technical term for shortenings used in the literature is 'subtractive derivation', which can include clippings and blends, or it can be applied to the deletion of phonological material which results in morphological change; this term is avoided because of its broader scope of reference. Besides, 'derivation' is a type of word-formation; labeling clipping a subtype of subtractive derivation implies commitment to a specific morphological taxonomy.

Whether clipping should be classified as a word-formation process is debatable. Much of the recent literature on clippings seeks to determine the patterns and constraints on clipping, treating it as a systematic process (e.g. Plag 2003; Lappe 2007), but the assumption is not universally shared. Haspelmath (2002: 25) excludes clippings from word-formation processes and chooses to view it as 'word-creation' because "the resulting new words do not show systematic meaning-sound resemblances of the sort that speakers would recognize." Clipping has also been subsumed under 'extra-grammatical morphology', a label employed by Dressler (1997) and Mattiello (2013), who discuss various taxonomic choices, including 'marginal' and 'expressive' morphology. No consensus has been reached. Durkin (2009: 116) considers the results of clipping unpredictable, and most recently, Don (2014: 27) places clippings outside of word-formation and writes that "clippings can result from deletion of just any part of the word, and there does not seem to be a clear pattern."

As the survey in Section 2 shows, "deletion of just any part of the word" is unnecessarily cavalier: the phonological and the morphological factors interact

---

1. "E. Partridge *Slang To-day & Yesterday* I. iii. 27: Slang delights to curtail (clip, abbreviate, shorten) words … Many such clippings have passed into standard English: *cab* for *cabriolet*, '*bus* for *omnibus*." (*OED*, s.v. *clipping*). Jespersen (1922: 170) uses "Stump-Words" as the title of the relevant section, but in the text itself he uses *clippings*. The OED attribution may need to be revised as to the first coinage of the term.

differently at different periods. The directionality of clipping in English changes quite dramatically after the 17th century. Thus, there are clear patterns emerging in English diachronically, and the trajectories of change call for an account that goes beyond randomness. Moreover, the differences in the treatment of this process historically suggest that it can be a window into the structure of the ambient synchronic systems. The macro-perspective on clipping allows an account in terms of well-formedness and faithfulness constraints: ONSET, STRESS, Morphological identity, Alignment (ALIGN-L). So, while I do not opine on what the most appropriate label for clippings should be – marginal, expressive, or extra-grammatical morphology – I hope that the diachronic perspective will broaden our understanding of the nature of the process.

The process of clipping is defined phonologically, morphosyntactically, semantically, and pragmatically:

(1)   Characteristics of clipping[2]
  - Prosodic: loss of at least one syllable
  - Morphosyntactic: no functional category shift
  - Semantic: typically preserves the denotative meaning
  - Pragmatic: typically involves shift of connotative meaning/register

'Clipping' as used here involves the loss of at least one syllable. Cases of consonant loss as in *know*, *soften*, *climb* fall under the more general terms of 'aphesis' or 'apheresis' for initial loss ([kn] > [n] in *know*), 'syncope' or word-internal cluster simplification ([ft] > [f] in *soften*), and 'apocope' or word-final cluster simplification ([mb] > [m] in *climb*). As these terms are also used to refer to syllabic loss, they are ambiguous with respect to the size and composition of the discarded material, the 'surplus'.[3]

---

2.   The characteristics in (1) can be extended to another productive process: the blending of two clipped forms: *chunnel* (1914) < *Channel tunnel; motel* (1925) < *motor hotel; spam* (1937) < *spiced ham*. These forms are not addressed because clipping is only one step in the process. I am also excluding suffixed or 'embellished' clippings, as in *movie* (1909), *commie* (1928), *commo* (1941); the unsuffixed clippings do not exist as free forms.

3.   *Surplus* for the discarded material and *residue* for the new form are terms suggested in Bauer & Huddleston (2002: 1634). Aphesis excludes loss of stressed syllables, which does occur in English, as in *shrooms* for *mushrooms, 'burbs* (1977) for *suburbs*.

The terminological survey is not exhaustive, see further Cannon (1989), Mattiello (2013). The term *aphesis* was originally an extension of *apheresis*, the latter involving specifically initial unstressed vowels as in *squire < esquire*, but now the two terms are used interchangeably (*OED*). Some sources, e.g. Pyles & Algeo (1993: 275), separate aphesis "a sound change" from clipping, so that *'féssor < professor* is aphesis (only unstressed syllables can be aphetic), but *prof < professor* is clipping, acknowledging that the results may 'look similar.'

Clipping does not affect the syntactic category of the input, though occasionally the clipped form's part of speech is that of the head of a phrase, while its phonological base is that of a modifier, as in *pub*, n. (1800) < *public house*; *perm*, n. (1927) < *permanent wave*.[4] Clipping can change the semantic content: *defence* (1297) > *fence* (1533) "the practice of using a sword"; *fanatic*, n. (1659) > *fan* (1889). It typically affects the pragmatic status of an item – clipped words "express primarily the attitudinal component of diminution, marking familiarity with the denoted object or concept. Like abbreviations, they often convey … in-group status … and are often restricted in usage to subgroups of the speech community" (Bauer et al. 2013: 402). This pragmatic aspect of clipping is recognized in all of the scholarship on the subject; it shifts the word's "linguistic value" and changes the "emotional background" (Marchand 1969: 441); it is a creative, sometimes ludic, process.

The usage characteristic in the last bullet in (1) is perhaps the least stable. With time, the downshift of register accompanying clipping may be reversed, as in *lunch* (1829) < *luncheon* (1580) "formerly objected to as vulgar" (*OED*), *mob* (1688) < *mobile* (*vulgus*) considered "miserably curtailed";[5] derogatory *miss* "whore" (1606) < *mistress* became a neutral title *miss* (1667) and *phone*, *plane*, *taxi*, *perk* are no longer just colloquialisms.

Traditionally, clippings are described in the relevant literature with reference to the original location of the rejected material: left-edge, middle, right-edge, or both edges (Marchand 1969: 441–450; Hansen et al. 1990: 148–150; Bauer & Huddleston 2002: 1634–1636; Mattiello 2013: 82; Miller 2014: 173–186). This linear concatenative taxonomy has held its value as an informative step in recent focused studies of clippings (see Berg 2011).

(2) Linear taxonomy of clippings
   a. Fore-clipping/left-edge: OF *engin* > *gin* (1200) "ingenuity"; *alone* > *lone* (1377); *defend* > *fend* (1330); *raccoon* > *coon* (1742); *suburbs* > *burbs* (1977)
   b. Middle/syncope: Lat. *anachōrēta* > OE *ancra* "anchoress"; OF *comencer* > ME †*comse*, v. "commence" (a 1225–1399); *sithen* > *sin* "thereafter" (1330); *fantasy* > *fancy* (1465)

---

4. Parenthesized dates after lexemes are the earliest recorded dates for that headword in the online *OED3* (http://www.oed.com/). The data in this study was collected using the 'Advanced Search' option in the OED for items labeled 'aphetic', 'aphesis', 'shortened', 'abbreviated', 'abbrev.' ('clipped/clipping' is used as an etymological term only 3 times) under 'Etymology'. The searches were further refined by 'Date of First Citation'.

5. "It is perhaps this Humour of speaking no more than we needs must which has so miserably curtailed some of our Words, … as in *mob. rep. pos. incog.* and the like." J. Addison in *The Spectator*, 1711, no. 135, cited in the *OED*, s.v. *mob*, n.[2] (and adv.)

   c.   Back-clipping/right edge: OF *baudetrot > bawd* (1362); Du. *genever* "juni-
        per" > *gin* (1689); *luncheon > lunch* (1829)
   d.   Fore- and back-clipping/ambi-clipping: *influenza > flu* (1839); *detective >
        tec* (1879); *refrigerator > fridge* (1926)

Rather than the location of the deleted material, clipping may also be described with
reference to the residue, or *retained* part. One taxonomy is based on whether the
residue is on the left or the right (Adams 1973; Bauer 1983). A theoretical extension
of this residue-centric analysis has been developed in studies of clippings in terms
of prosodic morphology, as in Nelson (2003) for French, Lappe (2007) for English,
Alber & Arndt-Lappe (2012), and Bauer et al. (2013: 402–404). These studies focus
on the *structure* of the output rather than on the location of the deleted material;
in this approach, the output conforms to some specified template, hence the label
'templatic morphology'.

   In (3), a taxonomy of clippings is presented according to the structure of the
output; what matters in the process is not what is taken away or where it is taken
from, but what part of the base is preserved.

   (3)   Clippings according to output

                        Monosyllabic: *coz* (1563) < *cousin; coon*
                        (1742) < *raccoon; doc* (1850, 1978)
                        < *doctor, or documentary*

      a.   Phonological

                        Disyllabic: *story* (1200) < *history; gator* (1844)
                        < *alligator, memo* (1705) < *memorandum*

      b.   Morphoprosodic:[6]
           tween (c. 1300) < *between*; *versal* (1599) < *universal*; †*panion*, n. (1553–
           1592) < *companion*; plane (1908) < *aeroplane*; *mare* (1994) < *nightmare*;
           klepto (1958) < *kleptomaniac*

The initial division here is based on the morphological status of the input: *coon*,
doc, *story*, *gator* in (3a) involve an input base which consists of a single root, while
in (3b) the input base has recognizable morphologically complex structure, as in
uni-versal, *night-mare*. In Lappe's (2007: 71) database of 702 word clippings, only
27, or 3.85%, belong to this morphologically complex type. Lappe distinguishes two

---

6.   This term is used in Lappe (2007: 70–71) for cases in which "the morphological structure of
the base word plays a more prominent role in determining the structure of the clipping." This is
an appropriate term for clippings involving identifiable and uniformly unaccented affixes, as in
Old and Middle English; see 2.1 and 2.2 below.

types of morphoprosodic shortening based on the residue: forms whose residue is a bound root or a combining form, such as *klepto*, and forms whose residue does not coincide with morphological boundaries, such as *prefab*, *illegit*. This latter type is arguably an extended case of phonological back-clipping as in (2c) and (2d). Importantly, she writes the following:

> The corpus used … does not contain enough data to allow any definite conclusions on morphoprosodic truncation. It will be the task of future research to further investigate this issue, empirically testing both the justification of the form class that has been set up there and the structural properties of these forms.
>
> (Lappe 2007: 71)

Since the early historical processes of clipping are heavily morphological, the current study attempts to address what Lappe saw as a 'future research task'. For a start, note that the linear typology in (2) and the residue-based typology in (3) reveal complementary yet different aspects of clipping; as we will see in Section 2, restricting the taxonomy obscures the diachronic picture. A macro-perspective on clippings in English requires reference to both directionality and the phonological and morphological properties of the residue.

Note the scope of coverage: The empirical data studied include clippings of common nouns, adjectives, adverbs, and verbs. Hypocoristic forms such as *Ed*, *Biz*, *Jen* are only referred to in passing. Also excluded are suffixed clippings (*comfy*, *hubby*, *combo*), reduced phrases (*will < last will*, *private < private soldier*), and clippings combined with blending and preserving the input compound information (*Caltech < California Institute of Technology*, *SoCal < Southern California*, *Tex-Mex < Texan and Mexican*; see fn. 2). Shortenings marked as 'graphic' in the *OED* are also ignored. From an analytic angle, the focus of this paper is *not* on why clippings occur in principle but why they happen in a particular way at a particular time in English.

The remainder of this paper is structured as follows. The next section presents a survey of the findings for Old English (OE) and Middle English (ME). Sections 3 and 4 address the analytical entailments of the diachronic data, compare these data to the current patterns of clipping, and explore the ways in which the long-term perspective enriches our reconstruction of English historical prosody and morphology.

2.   A historical overview of the data

While clipping in Present-day English (PDE) has been a much-studied topic in this century, as far as I can tell the diachronic perspective has not been revisited for nearly five decades.

Sundén (1904) and Slettengren (1912) are descriptive works with enduring value as a source of early data. Jespersen (1922, 1933, 1942) was a pioneer in moving beyond the classification of the available empirical material, treating it in terms of language acquisition and the drive towards monosyllabicity. Marchand's (1969) chapter on clippings, included "with reservations" (1969: 441) in his monumental book on English word-formation, became the standard reference and the basis of all subsequent diachronic comments (see Hansen et al. 1990: 146–150; Nevalainen 1999: 432–433; Kastovsky 2006: 213–214, 269). Unlike Jespersen, who considers the process one of the most characteristic traits of the language and uses it to support his general view of the superiority and monosyllabic efficiency of English (1942: 551), Marchand (1969: 441) believes that clipping "is not relevant to the linguistic system (*la langue*) itself, but to speech (*la parole*)." Both sources note the historical discrepancy between fore-clipping and back-clipping in earlier English and the PDE patterns, yet they do not take the discussion beyond that observation. Miller (2014: 173–176), the most comprehensive recent update of the research on the conditioning of various types of clippings, offers a classification with reference to the surplus location – right-edge (*mike ~ mic < microphone*) vs. left edge (*gator < alligator*) – as well as useful example dates from the *OED*, but he offers no diachronic analysis.

2.1   Old English

When digging into historical records, one has to recognize the conundrum created by the marked nature of shortened forms, which belong to an informal register of speech, and the type of formal written language that has survived. The impasse is real (i.e., it is likely that fewer instances of clipping are on record than actually occurred), though there are some side-trails that can at least confirm that linguistic shortcuts are no recent invention. *Grendel* is not *Grend and *Beowulf* is not *Wulfie in OE (though note Old High German [OHG] *Wolfo* for *Wolfbrand*, cited in Marchand 1969: 449), but a historical survey of hypocoristics shows some forms, as in (4):

(4)   OE hypocoristics
    Berter < Berhthere   (Bibire 1998: 165)
    Cutha < King *Cuthwulf*; *Sibba < Sigebeorht*   (Coates 2006: 325–326)
    Goda < Godgifu; *Saba <* (King) *Sæbeorht*; *Totta < Torhthelm* (Clark 1992: 459)
    OHG *Wolfo < Wolfbrand*   (Marchand 1969: 449)

Clark (1992: 459) writes the following: "The homilist Wulfstan's choice of *Lupus* as his pen-name might imply that familiarly he was called *Wulf*, but that is a

speculation." More intriguing is the discussion of the spelling of names on coins in Bibire (1998: 156), who points out that moneyers (die-cutters) were less well educated than the scribes; they were coming from a different social background, and therefore they were prone to use less conservative linguistic forms, or hypocoristic name-forms. He finds, interestingly, that there is a marked discrepancy between the treatment of king-names and moneyer-names on individual coins.[7]

With regard to non-hypocoristic clipping, a considerable set of alternate (i.e. clipped) forms that is fully documented consists of OE borrowings from Latin. First, we find fore-clipped forms that can be traced both to vulgar Latin, as is the case with *bishop*, and to parallel forms that are not attested in Latin, such as *pistelarie* < **epistelari* "the book from which the epistle is read". As Pogatscher (1888: 144) acknowledges, it is not always possible to determine whether the shortened forms arose in Germanic or whether they go back to colloquial Latin.

(5)   Clipping of Latin loan-words in OE
    a.   Fore-clippings
        OE *bisceop* < vulgar Lat. *(e)biscopus* < *episcopus*
        OE *stær ~ ster ~ steor* < *Lat. (h)istoria "story"*
        OE *magdala* < Lat. *amigdala* "almond tree"
        OE *moniaca* < Lat. *ammoniac*
        OE *pistle* < Lat. *epistola*; OE *pistelarie* < Lat. *epistolarium*
        OE *postol* < Lat. *apostol-*
        OE *spaldum* < Lat. *asphaltum*
        OE **spendan* (a-, for-) < Lat. *expendere*, or (later) OF *despendre*
        OE *Spene/Spenum* < (late) Lat. *Spānia* for earlier *Hispānia* (*Ispānia*)
        OE *fille* < Lat. *chaer(e)phylla "chervil"*
    b.   Back-clippings
        OE *cranic(e) ~ cranc* < Lat. *chronica* "chronicle"
        OE *fic ~ uic ~ fig* < Lat. *ficus* "fig-tree"
        OE *ylp*, ME *alp*, OE *elpend* < Lat. *elephantem*
        OE *kalend* < Lat. *kalendae, kalendas*, pl. "month"
        OE *regel ~ regl ~ reogol* < Lat. *regula* "rule, principle"
        OE *torr*, early ME *tūr* (<OF) < Lat. *turris "tower"*

---

7. Colman (2014: 131) finds Bibire's (1998) suggestion "untenable," but she offers no further comments on that point. She does, however recognize the presence in Old English of 'lall names' (2014: 126), "a category of words that belong to the nursery and are monothematic." She uses the term hypocoristic for shortenings based on non-monothematic names and offers an insightful analysis of lall-names originating in child language and subsequently adopted by the adults. On gemination in lall-names in Old English, see Colman (2014: 126–129, 146–150).

The variants involving clipping would have been familiar to at least some OE speakers; the significance of the examples in (5a) is that they provide a template for non-morphological clipping. The common tendency in (5a) is deletion of an initial unstressed vowel in nouns – there were no nouns in OE starting with unstressed <a-, e->. The new forms in (5a) have the properties defined in (1), yet they need to be qualified further: unlike post-ME clipping, all of the input words in (5) are loans and they violate the prosodic and phonological template for nouns in OE. The clipped variants can therefore be used as a diagnostic of the degree of nativization of the loans – cautiously – because of their uncertain provenance (see above).

A second set of clipped forms consists of back-clippings, exemplified in (5b); these clippings, as well, need further qualification. Pogatscher (1888: 148) labels the loss of material at the right edge *Suffixvertauschung*. Further examples of that pattern are shown in (6):

(6) *Suffixvertauschung* of loan-words

OE *abbod ~ ab(b)(e)* < Lat. *abbat-em* (acc.)	OE *camp* < Lat. *campus*
OE *altar ~ altare* < Lat. *altāre*	OE *casul* < Lat. *casula* "cloak"
OE *calc* < Lat. *calceus* "sandal"	OE *ceren* < Lat. *carenum* "wine"
OE *camel* < Lat. *camelus*	OE *cex* < Corn, Welsh *ceges*, "kex, keck"[8]
OE *mentel* < Lat *mantellum* "mantle"	OE *carr* < OWelsh *carrec* "rock"

The label in (6) is prompted by the loss of the original gender-marking case endings *-us*, *-a*, *-um* in the course of adopting Latin nouns. Not surprisingly, many of the shortened items are shared with other early Germanic languages. This is technically the same as back-clipping and probably signals colloquialization of the loan-word, but it would be a stretch to consider them precursors of the later patterns because the process is limited to loans and because it is not a typically 'insular' process. The usefulness of such forms, arguably, is that they attest to the availability, at least for the educated clerics, of doublet forms with identical left-edge alignment, the latter of increasing importance in later periods.

Another pattern of clipping widely attested in the native OE lexicon is illustrated in (7); the numbers are the counts given by *DOE*.[9]

---

8. The etymology is from the *Dictionary of Old English* (*DOE*: http://tapor.library.utoronto.ca/doe/); the *OED* takes the Welsh word to be a borrowing from English.

9. The current (2016) release of the *Dictionary of Old English* (*DOE*) has 15,533 headwords and covers approximately 27–30% of the total project, from A–H.

(7) Instability of *ge-* in OE
   a. OE fore-clippings as minority innovative forms
      bann (1) "command" < *gebann* (40)
      bǣru, pl. (2) "behavior" < *gebǣre* (55)
      bed (75) "prayer" < *gebed* (c. 1500)
      byrd (3) "birth" < gebyrd (110)
      būr (1) "free, but economically dependent peasant" < *gebūr* (24)
      dēfe (2) "fitting" < *gedēfe* (75); *dēfelic* (1) < *gedēfelic* (5); *dafenlic* (2) <
         gedafenlic (90)
      dwola (11) "error" < *gedwola* (140); *also dwol*, adj. (2) *"heretical"* <
         gedwol (6)
      fēa (2) *"joy"* < *gefēa* (c. 600)
      feoht (13) *"fight"* < *gefeoht* (c. 700)
      fēra (6) "companion" < *gefēra* (c. 450)
      hāt (4) "promise" < *gehāt* (125)
   b. OE *ge-* forms as minority innovative forms:
      bēacen "sign" (175) ~ *gebēacen* (1)
      drōf (6) "turbid, muddy" ~ *gedrōf* (2)
      dropa "drop" (70) ~ *gedropa* (2)
      fadung "arrangement" (15) ~ *gefadung* (10)
      fāh, *fāg* "hostile" (40) ~ *gefāh* (3)

In (7a), the token count is heavily in favor of the prefixed base, with the fore-clippings as a minority; then again, (7b) illustrates the inverse process, with the innovative addition of the prefix. The instability of *ge-* (along with that of *a-*) in OE, which is clearly attested in (7), is widely recorded and discussed. These prefixes were largely desemanticized (Minkova 2008); their increasing redundancy and interchangeability is a major contributing factor in their reduction and loss (see Hiltunen 1983). Among the six fully unstressable prefixes in OE (*a-*, *be-*, *ge-*, *for- of-*, *oþ-*), only *a-*, *be-*, and *ge-* are light syllables. All three can affect the syntactic categorization, i.e. they are head-prefixes, but the headhood of *a-* and *ge-* became progressively obscure, starting in the north in the 10th century (Slettengren 1912: 101–102; Hiltunen 1983), a development associated with Scandinavian influence.[10]

The significance of the variants in (7) goes beyond their grammatical and semantic nature. More relevant for the history of clipping is the existence of native doublet forms, i.e. with or without a left-edge light unstressed syllable that is only marginally identifiable as a morpheme, whereby *ge-* is less identifiable than *a-*. A remarkable, and previously unnoticed, fact about the examples in (7a) and (7b) is the striking

---

10. Hiltunen (1983: 94–102) makes a strong case for 'semantic multiplicity' as a trigger of the breakdown of prefixation in OE; see similarly Kastovsky (1992: 377).

contrast in the frequency of attestation: the clipped forms in (7a) are the minority, while the reverse is true of the prefixal forms with an innovative *ge-*. The 'decapitated' forms (Slettengren 1912: 102) in (7a) are incipient, low-frequency variants, and it is likely that dropping the *ge-* at the left edge correlates with colloquial usage. Phonologically, the variation in (7) is the same as that of the loan-words in (5a).

   At this stage, i.e. late OE, the dominant pattern of shortening native words was fore-clipping, a pattern that becomes marginalized in PDE (see (14) in Section 3.4). At the right edge, the fact that at least *some* hypocoristic forms, as in (4), have made it into the records suggests, in parallel with current usage, that there is no diachronically traceable change in the way speakers treat names; the process is typologically the same, but the special register of such forms has kept them out of the earliest records, though see Colman (2014) for extensive detective and analytical work in that area.

2.2   Middle English

Turning to ME, we find the pattern in (5a) and (7a) recurring in loans from Anglo-Norman (AN) and Old French (OF).

   (8)   Mixed phonological and morphoprosodic fore-clipping in AN and OF loans[11]

mend, v. (1225) < *amend*	*spute* (1225) < *dispute*
cheker, n. (1330) < *escheker*	*stress* (1303) < *distress*
cumber (1303) < *encumber*	*scomfit* (1303) < *discomfit*
fend, v. (1330) < *defend*	*liver* (1330) < *deliver*
stoun(d), v. (1330) < *estoner*	*senye* (OE, 1225–) < *ensign*
quest (1350) < *request*	*liverance* (1380) < *deliverance*
cense (1382) < *incense*	*minish* (1382) < *diminish*
monishing (1382) < *admonishing*	*colet* (1382) < *acolyte*
print, v. (1382) < *imprint*	*faunt* (1382) < *infant*
bandon (1400) < *abandon*	*till* (1398) < *lentil*

The majority of the attestations in (8) are instances of dropping of a Romance prefix (*a-*, *ad-*, *de-*, *dis-*, *en-*, *in-*), not unlike the cases in (5a) and (7a). The set as a whole is labeled 'mixed', because the discarded material is not necessarily an identifiable

---

11. The bundling together of Anglo-Norman and Old French is a shortcut: separating the sources, as Slettengren (1912) does, reveals an important distinction: the rate of fore-clipping in Anglo-Norman is considerably higher than the corresponding rate in Old French; see Pope (1934: 439) on the semantic bleaching of prefixes and the higher rate of pretonic syllable loss in Anglo-Norman.

prefix, even in the donor language, e.g. ME *cheker* < AN *escheker*, OF *eschekier* "chess"; ME *scomfit* < AN *descounfit, descumfit*, OF *desconfit*; ME *spute* < OF *desputer*; *faunt* < OF *enfaunt, enfant* "infant"; ME *till* < *lentil*. These fore-clippings are homegrown.

As a whole, the ME set in (8) reinforces patterns existing before, where reduced productivity of native prefixation makes word-initial iambs # σ ɕ (-) # rarer, boosting the model of monosyllables (*fend, quest, print*) or of trochees # ɕ σ # (*checker, cumber, minish)*. The outputs have left-aligned morphological stress, the dominant prosodic contour of the Germanic component of the vocabulary. Not surprisingly, then, we find that the OE template in (7a) continues to be robust in the native lexicon of ME.

(9)   Native fore-clipped forms in ME

back, adv. (1300) < *aback*   *bout* "about" (1250) < *abuten*
mang (?c. 1200) < *gemong, onmang*   *lone* "alone" (1377) < *alone*
ȝæn (?c. 1200) < *onnȝæn*   *slant*, adv. (1495) < *aslant*

The forms in (7a), (8), and (9) share yet another property: the initial syllables of the output words in all three sets have a filled *onset*. This is an obligatory well-formedness condition on all fore-clippings in English. I am not aware of any instances of fore-clipping resulting in a vowel-initial word in OE or ME, whether the output is monosyllabic or not. Since fore-clipping is a minority pattern in PDE, as are onsetless words, it is to be expected that onsetless fore-clippings would be rare, yet among the post-ME fore-clippings we do find, for instance, *atomy* "skeleton" (1597) < *anatomy*; †*omy*, adj. (1674) < *loamy*; *oodles* (a1867) < *scadoodles* ("perhaps" the right etymology; see *OED*); †*outang* (1869) < *orangutan*; *oscine* (1892), a chemical term borrowed from German < *hyoscine*; the biochemical term *opsin* (1951) < *rhodopsin*. This is a small but notable difference between fore-clipping earlier and now, accompanying the much more significant switch from fore- to back-clipping in PDE, where *ad, amp, ap, info, op, ump* are unexceptional.

The avoidance of vowel-initial stressed words in earlier fore-clipping is important for the overall history of the ONSET constraint in English: it was inviolable for stressed syllables in Old English, and as argued in Minkova (2003) the constraint was gradually demoted in ME. The data on fore-clippings are a window into the continuing relevance of ONSET into PDE; moreover, the sensitivity of the output to syllable structure underscores the importance of including directionality in the analysis of clippings.

The distinction between morphoprosodic fore-clipping, where the surplus is an identifiable productive prefix (OE *gebed* > *bed* "prayer"; ME *defend* > *fend*), and fore-clipping, where the morphological autonomy of both the surplus and the residue is unclear (ME *i-fihte* > *fight* vs. *distress* > *stress*), became increasingly opaque

in ME. This initiated a shift in the direction of purely phonological fore-clipping. This is illustrated by the generation, and sometimes quite prolonged use, of new forms such as in (10), where the parenthesized dates mark the chronological span of their attestation:

(10)   Non-morphological fore-clipping in 'archaic' entries in the *OED*
   † *sturb*, v. (c. 1225–1450) < *disturb*
   † *vangelist*, n. (1330–1567) < *evangelist*, also *vangel*, n.
   † *spittle*, n. < *hospital* (1225–1839); also in Dutch and German
   † *rabite*, n. "an Arabian horse" (1330–1500) < *Arabit* (adjective)
   † *gin*, n. "a mechanical device, machine" (c1386–1621) < OF *engin*[12]
   † *till*, n. (1398–1760) < *lentil*, n.
   † *sparagus, n.* (1543–1668) < *asparagus*[13]

With the exception of OE hypocoristics as in (4),[14] the few items in (5b) and the more general *Suffixvertauschung* of loan-words in (6), there have been no cases of back-clipping in OE comparable to the patterns observed today. This situation continues in ME.[15] The closest ME parallels to back-clipping involve some stems which originally end in [-ən ~ n̩]#: OE *eln ~ elin*, ME *eln(e)* "ell"; OE *gamen* > ME *game* "game"; OE ***mæden, mægden*** > ME *mæide* "maid"; similarly PDE *eve, morrow*. The obstacle to including them in the overall history of back-clipping is that the narrowly defined deletable portion in them is identical to inflectional <-en> in plurals, infinitives, strong past participles and past indicative plurals; the shortening in those few items is most likely analogical.

---

12. As one of my reviewers remarked usefully, "the same spelling gin can indicate either a fore- or a back-clipping. This increases the ambiguity of abbreviated terms (vs. the full forms)." The form *gin* < *engine* in other senses predates and postdates *gin* < Du. *genever* "juniper". It is not clear, however, that the ambiguity was a problem for 19th-century users, when the timeline of quotation evidence in the *OED* peaks for both words.

13. The *OED* has a separate entry for †*sparagus* with the date-span cited here, but it also has '*sparagus* entries under 'asparagus' as late as 1785.

14. Hypocoristic forms in ME are much more broadly attested; "baptismal names were used in a wide variety of hypocoristic or pet forms, especially by ordinary folk"; see McClure (1998). He cites, among others, the names of the rioting peasants of the 1381 poll tax revolt epitomized as *Walle, Thomme, Symme, Belle, Gibbe, Hykke, Colle, Geffe, Wille, Grigge, Dawe, Hobbe, Lorkyn, Hudde, Judde, Tebbe, Jakke,* and *Hogge*.

15. I am not including the loss of stem-final /-ə/# or various inflectional endings in ME, although these processes have the first three of the features of clippings defined in (1) and possibly correlate with register of usage. Loss of stem-final /-ə/# differs from PDE back-clippings in two very important ways: its diffusion was primarily morphosyntactically and not pragmatically driven, and it was tightly linked to language contact.

2.3   Later patterns of clipping

Both morphoprosodic and purely phonological fore-clippings continued to be formed after the 16th century, but at a much slower rate.

(11)   Post-1500 fore-clippings[16]

a. *Morphoprosodic fore-clippings*
tain, v. (1501) < *obtain*
judicate, v. (?1577) < *adjudicate*
trusion (1604) < *intrusion*
versity (1680) < *university*
pike (1812) < *turnpike*
roach (1822) < *cockroach*
bus (1832) < *omnibus*[17]
phone (1880) < *telephone*
plane (1908) < *airplane*
chute (1920) < *parachute*
tingle (1930) < *whelk-tingle*
fax (1946) < *telefax(imile)*
hood (1969) < *neighborhood*

b. *Phonological fore-clippings*
gypsy (1537) < *Egýptian*, n.
rebesk, adj. (1559–1656) < *Arabésque*
possum (1613) < *opóssum*
wig (1675) < *périwig*
squash (1678) < *músquash*
coon (1742) < *raccóon*
roo (1841) < *kàngaróo*
gator (1844) < *àlligátor*
tash (1894) < *moustáche*
toon (1932) < *cartóon*
roid (1978) < *stéroidjamas* (1960) < *pyjámas* also *jammies* (1928)
sheen (1960) < *machíne*

Fore-clipping after 1500 shows more of the transitional characteristics already in evidence in (8). The patterns in (11a) are an extension of the earlier patterns shown in (7a), (8), and (9). In addition to the familiar dropping of monosyllabic prefixes as in *tain*, v. (1501) < *obtain*, *trusion* (1604) < *intrusion*, the morphoprosodic set of left-edge clippings show innovative features: deletion of more than a single syllable, e.g. *versity* (1680) < *university*, and deletion of a morpheme bearing primary stress, as in *pike* < *turnpike*. Often the product of fore-clipping is based on secondary stress, as in *roach, bus, phone, plane, hood.*

The patterns in (11b) present a practically exhaustive list of post-1500 fore-clippings that do not involve a recognizable morphological component either in the reject or in the new form. In terms of prosody, *for the first time* phonological fore-clipping is not categorically bound to lack of stress in the deleted portion. The surplus syllables in the shaded examples in (11b) – *wig, squash, roo, gator, roid* – bear primary stress, though in some cases the residue reflects the secondary stress of the base form, as in *alligator* and *kangaroo.*

---

16. 1500 is admittedly an arbitrary date, but if we want to draw a timeline, it seems that 1500–1800 is the period of lowest productivity for clippings as recorded in the *OED*.

17. The unit *omni-* is identified as a combining form, recorded since 1593; the <-i> is from the dative plural *-ibus.*

The mixed picture at the left edge is complicated further by the potential of resyllabification at word boundaries when the word mimics a clitic group: an unstressed initial vowel is dropped but the consonant remains in the new form, creating a well-formed output with a filled onset, as exemplified in (12):

(12)   Fore-clipping in conjunction with clitic group misanalysis
   leven (a1400) ~ *eleven* ~ *en(d)leve(n)* "eleven"
   noint (1400) ~ *anoint*
   nentes, prep. adv. "in line" (1400) < ME *anentist*
   larum (1453) "alarm" < Middle French *<à l'arme>* ~ *alarme*
   nemel (1440) ~ *enamel*
   notomy (1487–1994) ~ *atomy* "skeleton" (1597–1995) < *anatomy*
   emony (1644–1882) < *anemony*
   notty (1725) < *anatta* (AmInd "orange red dye")
   Also *treat* "treatise"(1400–1555) < *tretis, treatise*, n.; *beck*, v. (1300) < *beckon*[18]

In the foregoing, it has been seen that fore-clipping is the dominant type of clipping in OE–ME. In its earliest manifestations, the discarded material is limited to unstressed monosyllables. In the native ME wordstock, the loss is morphoprosodic, but in the loan vocabulary the morphological constituency of the surplus is often obscured. Along with that, the prosodic association between lack of stress and deletability is also destabilized. As for the numbers, my search of over 700 forms identified as aphetic ('fore-clipping') in the *OED* shows that over 600 are recorded before the beginning of the 17th century, and about 96% of them are recorded before c. 1800.

All surveys of clipping in English note (i) the lateness of back-clipping and (ii) the preponderance of back- over fore-clipping in PDE.

(13)   Back-clipping 1500–1700

hake (1538) < *haquebut* "firearm"	*prop* (1607) < *proposition*
coz (1563) < *cousin*	*brandy* (1640) < *brandwine* ~ *brandewine*
con (1572) < *contra*	*tymp* (1650) < *tympan*
buck (1577) < *buckwheat*	*fan* (1682) < *fanatic*
chap (1577) < *chapman*	*chum* (1684) < *chamber-fellow*
†*scull* "servant" (1566–1743) < *scullion*	*hack* (1687) < *hackney*
sol (1588) < *solution*	*mob* (1688) < *mobile (vulgus)*
lunch (1591) < *luncheon*	*incog* (1699) < *incognito/a*
miss (1606) < *mistress*	*strum* (1699) < *strumpet*

---

18. With regard to *treat*, the *OED* (s.v. *treat*, n.¹, †8) states: "? A clipped form of *tretis, treatise*, n., the *-is* being taken as plural suffix." With regard to beck, "the *-en* … [is] apparently taken as the infinitive ending" (*OED*, s.v. *beck*, v.).

Prosodically, the surplus portion of the forms in (13) is mostly unstressed, though output forms such as *sol < solution, prop < proposition, fan < fanatic* show a preference for alignment with the left edge of the word over stress preservation. Crucially, the new forms are different from the examples of *Suffixvertauschung* of loan-words shown in OE, as in (6). Back-clipping was not a demonstrably productive process in ME, and it is very sparsely attested between 1500–1700. Then, in the late 18th century, the trajectory of new forms in this group goes up steeply. Figure 1 shows the historical distribution of forms – the data on left-edge clippings are my counts extracted from the *OED*, while the numbers on back-clipping are those cited in Lappe (2007: 79–70); comparable though not exactly the same counts are found in Davy (2000) and Berg (2011).
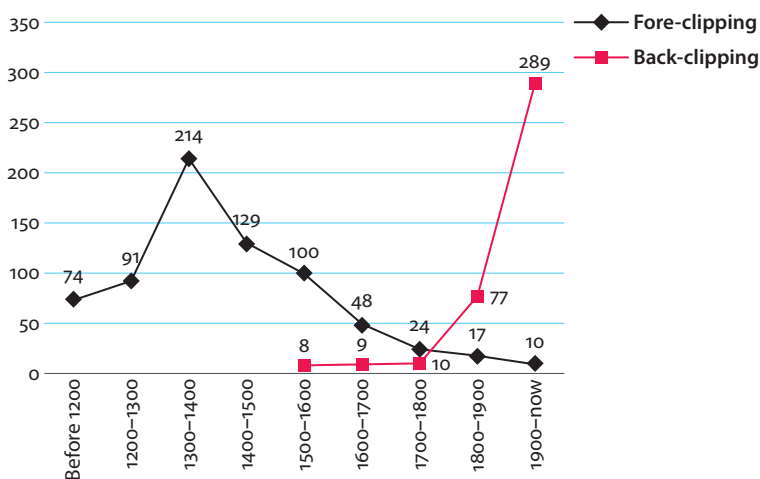
Figure 1. Fore-clipping vs. back-clipping from OE to PDE

Upwards of 75% of the words the *OED* identifies as 'shortened' are instances of back-clipping and their first entries are recorded after 1800. Again, although studies of clipping in English invariably note that the dominant model of clipping in PDE is back-clipping, the quite dramatic diachronic reversal of the directionality of this process has not been addressed in the literature.

3. The big picture: Why does it matter?

According to many sources (Pyles & Algeo 1993: 285; Algeo 1998: 84–85; Davy 2000: 60), shortening processes are increasingly productive in English, generating 9–15% of new words. Algeo (1998: 84–85) estimates that as many as 17.5% of the

post-1900 new words in English are shortenings, but his umbrella term includes acronyms, initialisms, and back-formations. In any case, the portion of the vocabulary to which clippings belong is growing rapidly and provides new analytical material. This has not escaped linguists' attention: the properties of clippings and other shortenings in PDE have been the target of numerous investigations.

As already noted, however, the data behind the striking historical asymmetry in Figure 1 has never been explicitly described or contextualized. Covering every aspect of the diachronic picture is beyond the scope of this study, but we can start by addressing the evolution of the modes of clipping with respect to four specific phonological and morphological features: ONSET, STRESS, Morphological identity, and ALIGN-L.

ONSET is a prosodic markedness constraint: syllables must have onsets (see Kager 1999: 93). STRESS is a shortcut for a faithfulness constraint that ensures that the stress of the residue mirrors the stress of the input form; it differs from Lappe's (2007: 192) MAX-$\sigma_{1+\text{stress}}$ constraint which requires that "every segment that is in the initial syllable and has some degree of stress in the base form has a correspondent in the truncated form" – it bundles together stress and initial position, which is inexpedient for the diachronic material. 'Morphological identity' is prose for the faithfulness constraint MAX-BT (Kager 1999: 264), which ensures that every (morphological) element in the input string (B) has a correspondent in the residue (truncated form, T). In the case of historical clippings it is the root that satisfies the constraint, while prefixal fore-clipping violates it.[19] ALIGN-L as used here stands for ALIGN-L (WORD) = ALIGN-L (GRW) for 'grammatical word' (Kager 1999: 118), which matches the left edges of the base and the residue. Lappe's (2007: 176) positional constraint MAX-σ_1, "every segment in the initial syllable in the base form has a correspondent in the initial syllable of the truncated form" does the same job without reference to morphological structure.[20]

3.1   Clippings as new evidence for ONSET

One of the findings from the survey in Section 2 was that historical fore-clipping invariably respects the ONSET constraint; see (7a), (8), (9). It is only towards the end of the 16th century that we get some isolated instances of fore-clipping resulting

---

19. A detailed formal account will require the separation of the two constraints.

20. I avoid OT constraint formalization in tableaux, though the proposed changes in the relative importance of these factors could be tested and presented formally, both in OT and in other models – the goal here is to draw attention to a previously unmined set of diachronic material.

in a Vowel-initial (V-initial) residue. The early history of this constraint, a very important parameter in the description of PDE syllable structure, has so far been reconstructed primarily on the basis of metrical usage (Minkova 2003) and clitic misanalysis, as in (12). The hypothesis based on that evidence is that ONSET was inviolable in the native OE vocabulary and continued strong in ME. Looking at the history of early clippings has yielded another piece of evidence in favor of this constraint's persistence in ME. Moreover, the list of PDE clippings in Mattiello (2013: 287–295) does not show a single vowel-initial fore-clipping, including inputs with transparent morphemic structure such as *burb* (1997) < *suburb*. Recent studies of the fine-grained segmental constraints on clipping in modern English empha-size the importance of the syllable coda (only 4.8% of clippings are coda-less; see Lappe 2007: 140), but for the historical linguist the persistence of ONSET is equally important. The rarity of V-initial fore-clippings is doubly motivated: fore-clipping as a productive pattern has been receding, and although the general vocabulary allows V-words such as *able*, *eddy*, *image*, *oral*, such words remain a statistical mi-nority.[21] The innovative type of clipping, namely back-clipping, prioritizes ALIGN-L and also tolerates V-initial forms.[22] The interplay between lexical onsetless forms and onsetless prefixation (e.g. *in-*, *en-*, *a-*, *of-*, *up-*, *-un-*) in the history of clippings in English is a matter for further research.

3.2   Clippings and the salience of stress

In OE, fore-clipping was well attested (see Section 2.1), both in loan-words and in the native lexicon, with the generation of new forms obeying the conjunction of STRESS and ONSET. Fore-clippings of loan-words, as in OE *pistle* "epistle", *postol* "apostle", produce words that fit the Germanic structural and prosodic template, which is characterized by left-aligned main stress and filled onset. The instability of OE *ge-* in native words can be attributed to the same factors: eliminate initial unstressed syllables, main-stress to the left, and filled onset. In native as well as in loan-words, the grammatical and semantic content of the discarded part is ob-scured. The opacity of the morphological composition of native OE *ge-*, ME *y-* input words makes the Morphological identity constraint violable.

---

21. I have no data on the ratio of CV- vs. V-lexical monosyllables in English, but only 9% of trochaic disyllables in PDE are of the type *able*, *eddy* (see Minkova 2003: 176).

22. Only 8.8% of the PDE clippings (Lappe 2007: 140) are of the type *ump* < *umpire*. They are all back-clippings. In PDE name truncation, a filled onset in the initial syllable also makes the preservation of that syllable more likely (Lappe 2007: 105–106).

In Middle English, fore-clipping continued in the native word-stock and was strongly reinforced by the adoption and adaptation of numerous AN and OF loan-words. ME fore-clipping is enabled primarily by lack of stress at the left edge of the word, producing *lone < alone, slant < aslant, stress < distress, fend < defend*, etc. The rejection of the original iambs at the left edge attests to the continuing strength and stability of the Germanic prosodic contours: STRESS and ONSET continue to be the dominant factors in shaping the residue. Further, even though some fore-clippings in (7) were attested in Old French, the majority of them entered the language from Anglo-Norman. This is another new angle on the ways in which Anglo-Norman diverged from Continental French (Pope 1934: 437–439). Whereas the majority of the ME fore-clippings are instances of prefixal deletion, we also get forms such as *spute, vangelist, rabite*, where the process is purely phonological and neither the surplus nor the residue is based on morphologically autonomous elements.

Post-1500 fore-clippings such as *tain, trusion, versity, gypsy, possum, coon* also obey STRESS, but its effect is gradually demoted, such that by the 17th century the surplus can be a syllable with primary stress: *(peri)wig, (mus)squash, (cock)roach, (omni)bus*. Situating this gradual downward shift of the perceptual importance of STRESS within the overall history of English prosody is instructive: by 1500, the Germanic stress rule had become threatened by the competing right-to-left stress assignment (see Minkova 1997; 2014: 306–310). One can take this a step further and suggest that at least one of the factors behind the striking reversal of the directionality of clipping after the 17th century is related to the hybrid nature of stress in PDE. The availability of multiple stress contours for the loan vocabulary (i.e. #σ ɔ́ σ#, #ɔ́ σ σ#, #ɔ́ σ ɔ̀#, #ɔ̀ σ ɔ́# for tri-syllabic words and a similar proliferation of patterns for longer words) make it more likely that the Germanic #ɔ́ (σ)# prosodic template for unaffixed words can be achieved by dropping the right edge, by back-clipping.

As back-clipping implies preservation of the word's left *edge*, the main innovation identifiable in word-clipping in PDE is ALIGN-L. Preservation of the left edge in PDE is the dominant pattern of clipping (84.9% in Berg 2011: 7; see also Lappe 2007: 152, 181), irrespective of whether the word-initial syllable is stressed or not.[23] While the conjoint effect of STRESS and ALIGN-L obviously enhances the probability of back-clipping, even in words with non-word-initial stress, like *legitimate*, back-clipping is as high as 91.9%.

---

23. Carter & Clopper (2002) confirm the importance of ALIGN-L for adults but not for children, suggesting that "there seems to be a shift in word edge bias at some point in time from right (word-final) to left (word-initial)."

(14)   Effect of stress on back-clipping vs. fore-clipping in PDE[24]

Base form   Examples

Back-clipping 95.3%   *business > biz*

Initial stress:

Fore-clipping 4.7%   *suburbs > burbs*

Back-clipping 91.9%   *legitimate > legit*

Non-initial stress:

Fore-clipping 8.1%   *machine > sheen*

3.3   Morphological identity and clipping

The morphological identity of the residue in OE depended on the stratification of the vocabulary. Fore-clipping in loan-words, e.g. OE *pistle* "epistle", *postol* "apostol", does not occur at a transparent morphological boundary; dropping the initial unstressed syllable is part of the process of prosodic nativization. In the native vocabulary the semantic and functional bleaching of consistently unstressed light-syllable prefixes obscures their morphological nature, e.g. *delf ~ gedelf* "ditch", *fera ~ gefera* "companion". However, the morphological compositionality of the input remains important in that clipping did not affect the shape of the root.

In AN and OF borrowings, the morphological constituency of the input would have been non-transparent to a native speaker. The break of the link between prefixal status and stress makes lack of stress the main enabling factor for clipping, e.g. *spute*, *mend*, *stress*. The residue is no longer coextensive with identifiable morphological material.

Today the involvement of morphology in clipping remains a vexed issue. In back-clipping with monosyllabic residues, neither the surplus nor the residue can be identified as a morpheme: *croc[odile]*, *lab[oratory]*, *ump[ire]*. In non-monosyllabic back-clippings, the residue can be modeled on previously existing and recyclable 'combining forms', the term used by the *OED* for bound classical morphemes used in compounding: *pseudo-* (1425), *hypo-*(1450), *proto-* (1576), *pyro-* (1593),

---

24. These percentages are from Berg (2011: 8).

gyneco- (1594), *Gallo-* (1601). In those early combining forms, the *-o* is directly inherited from Greek 'connective' *-o-* in compounding, which was retained and was independently productive in post-classical Latin. Back-clippings based on this 'connective' *-o* are exemplified by *memo* (1705), *hypo* (1711), *demo* (1794) < *democrat*, *loco* (1833) < *locomotive*. It may be the density of *-o* combining forms and *-o* back-clippings that have led to the emergence of *-o* as an independently productive suffix, as in *wino, cheapo, whammo, commo*.[25] Sixty-eight out of ninety-nine, or 68.7% of the *-o* data in Lappe (2007: 69–70), inherit *-o* from the input base. For this subset of back-clippings, therefore, morphological faithfulness remains a factor, albeit low-ranked and violable, and definitely phonologically restricted.

Fore-clipping presents a somewhat stronger case for the continuing relevance of the morphological structure of the input. A considerable portion of the fore-clippings start out as word-formation products: *plane, roach, hood, phone, pike* are based on compounds or compound-like forms with a prosodic contour of primary–secondary stress. Whether such fore-clippings, arguably based on secondary stress, are illuminating with respect to the distinction between compounds and noun phrases is a matter which I leave open. Thus, as most recently remarked by Miller (2014: 183), fore-clipping in PDE remains an analytical challenge within current accounts.

3.4   Clippings and left-alignment

The dominance of back-clipping in PDE is well documented. This innovative pattern has been extensively analyzed – it preserves the left edge of the input. As far as we can tell from the existing records, alignment to the left word-boundary played no role in OE, unless one considers the general Germanic *Suffixvertauschung* and an isolated form such as OE *ylp ~ elpend*, ME *alp* < Latin *elephantem* distant harbingers of later changes. The observance of left-edge alignment grew steadily after the 15th century aided by the rapid growth of the loan vocabulary and the prestige of classical learning. One of the diagnostics of its rising importance is the restoration of the input left-edge material in relatively short-lived forms such as in (9): †*sturb*, †*vangelist*, †*spittle*, were abandoned in favor of *disturb, evangelist, hospital*.

---

25. The *OED* editors suggest that back-clippings such as *memo* and *hypo* "probably established an association of the ending *-o* with casual or light-hearted use which it has retained ever since." They also note that "After 1851 this type of clipping becomes, and has remained, extremely common."

4. Summary: The shifting priorities in clipping

The only recorded type of back-clipping in OE is that of first names; hypocoristic shortening is also recorded in ME and strongly productive in PDE. That type of truncation is therefore typologically stable.

The typology of common-noun clipping has undergone important shifts. The textual records of OE and ME do not show back-clipping; it started in the 16th century, possibly on the model of hypocoristic clipping. At that point, the first syllable of the word, stressed or unstressed, became the anchor of the new form. This is manifested both in the gradual decrease of fore-clipping and in the rise of iambic forms of the type *exec*, *celeb* along with the more predictable *fab*, *temp*. The violation of non-finality, a prosodic constraint which disallows primary stress on final syllables in OE and which became more widespread in late ME (Minkova 1997), accounts for the broadening of the available output options. The new hybrid model of prosodic parsing in PDE – partly morphological, partly Germanic, partly from right-to-left – is reflected in the outcome of clipping.

The historical shift of the phonological and morphological factors identified in this survey is shown in (15):

(15) Continuity and innovation in English clippings

OE   STRESS, ONSET >> Morphological identity >> ALIGN-L
ME   STRESS, ONSET >> Morphological identity, ALIGN-L

　　　　　　　Fore-clipping ONSET >> STRESS >> Morphological identity

PDE

　　　　　　　Back-clipping ALIGN-L >> STRESS >> ONSET >> Morphological identity
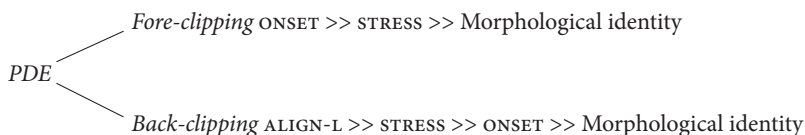
What we have gained from the diachronic perspective is, among other things, a better understanding of the difference in the directionality of clipping today. Clippings confirm the long-known tendency towards monosyllabicity (Jespersen 1933). For PDE, Berg (2011) proposes an account in which the choice of discarding front matter versus discarding material at the right edge can be seen as a matter of production favoring fore-clipping vs. perception favoring back-clipping. Back-clipping causes less perceptual disruption. This is related also to the pragmatic-communicative aspect of name-clipping vs. common-noun clipping: "first names are ordinarily heard in highly restrictive contexts such as family circles" (Berg 2011: 12), which allows a higher rate of fore-clippings for such items.[26] These are certainly valid pragmatic

---

26. In Berg's account, fore-clipping is more susceptible to intentional or unintentional errors, 'slurring' and 'mutilating' in production, and therefore it is more easily tolerated in restricted contexts, i.e. for personal names. Berg shows that the rate of deletion of stressed initial syllables is

attributes of the *existing* imbalance between fore- and back-clipping, but they do not address the emergence of back-clipping at a particular diachronic stage, since they are universal and not period-specific.

Examining the history of clippings in the context of the properties of the ambient language at earlier stages opens up additional explanatory strategies. In Early Modern English as many as 1500 new loans enter the lexicon during each decade between 1500–1700 (Minkova & Stockwell 2009: 47). This affects the diversity of available prosodic templates. Main stress is no longer a necessary property of the first root syllable, nor is lack of stress inviolable in prefixation: Chaucer has initial stress on the nouns *subject*, *proverb*, *perfect* (Minkova 2014: 311). As stress is no longer the primary word-boundary marker, that function begins to be shared with any left-edge syllable. The entailment for clipping is increased occurrence of back-clipping. At the same time, fore-clipping, traditionally targeting unstressed prefixal material, is counteracted by broader familiarity with the source language and restoration of the original lexeme, as in *monish* > *admonish*, *minish* > *diminish*, *faunt* > *infant*. It is also the case that the attestations of fore-clipping carry more social stigma as being 'lazy' (see fn. 26 on the association of production and fore-clipping), while PDE back-clippings – *pash* (1891), *artic* (1945), or *tranq* (1967), etc. – are simply 'cute', 'playful', and 'in-group' but lack the stigma of ignorance or carelessness.

Re-examining clippings diachronically enriches the history of ONSET in English. The fact that ONSET remains inviolable in fore-clipping from OE to ME and shows a continuing presence in PDE is a new window into the reconstructed history of that constraint. While back-clipping is fully congruent with a top-ranking left-alignment in PDE, fore-clipping mostly continues the historical template of high-ranked STRESS. Recoverability of the morphological identity of the base components is a viable option in fore-clipping.

Seen from a broader analytical point of view, the different treatment of ONSET in clippings of different historical periods and the interplay between different prosodic systems and the type of clipping strengthen the position that not all clipping is aleatory, or completely unregulated. The historical perspective adds another angle to studies claiming that clipping is a non-arbitrary process (see Fischer 1998: 41–43; Lappe 2007; Jamet 2009; Berg 2011; Alber & Arndt-Lappe 2012; Mattiello 2013; cf. the reappraisal in Miller 2014: 182–184).[27]

---

higher for names (12.2%) than for common-noun clippings (4.7%). The difference is statistically significant.

27. Since the first version of this paper was presented in July 2014, it does not include references to post-ICEHL 18 work addressing the structural properties of clippings: Ronneberger-Sibold (2015), Spradlin & Jones (2016), Arndt-Lappe (2018). These works do not address the diachronic data for English presented here.

Further research can go in many directions. Can clippings be used as evidence for the rate of nativization of loan-words in earlier English? Are vowel-initial back-clippings attested at the same rate as vowel-initial words in the ambient lexicon? What is the effect of frequency of the prefix and the frequency of the root of the base forms on the occurrence and survival of clippings? Are there phonological factors that favor, or even guarantee, the longevity of early fore-clippings – why did *mend*, *fend*, *stress*, *bus* survive while so many others were short-lived? Can clipped forms be diagnostic of relaxed social attitudes in earlier texts? Can history help us speculate on the future of clippings?

Abbreviations

acc.	accusative	OF	Old French
adj.	adjective	OHG	Old High German
adv.	adverb	Lat.	Latin
AmInd	American Indian	n.	noun
AN	Anglo-Norman	PDE	Present-day English
Du.	Dutch	pl.	plural
ME	Middle English	prep.	preposition
OE	Old English	v.	verb

References

Adams, Valerie. 1973. *Introduction to modern English word-formation*. London: Longman.

Alber, Birgit & Sabine Arndt-Lappe. 2012. Templatic and subtractive truncation. In Jochen Trommer (ed.), *The phonology and morphology of exponence*, 289–325. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199573721.003.0009

Algeo, John. 1998. Vocabulary. In Suzanne Romaine (ed.), *The Cambridge history of the English language*, vol. 4, 57–91. Cambridge: Cambridge University Press.

Arndt-Lappe, Sabine. 2018. Expanding the lexicon by truncation: Variability, recoverability, and productivity. In Sabine Arndt-Lappe, Angelika Braun, Claudine Moulin & Esme Winter-Froemel (eds.), *Expanding the lexicon: Linguistic innovation, morphological productivity, and ludicity*, 141–170. Berlin: De Gruyter Mouton.

Bauer, Laurie. 1983. *English word-formation*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139165846

Bauer, Laurie & Rodney Huddleston. 2002. Lexical word-formation. In Rodney Huddleston & Geoffrey K. Pullum (eds.), *The Cambridge grammar of the English language*, 1621–1723. Cambridge: Cambridge University Press.

Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198747062.001.0001

Berg, Thomas. 2011. The clipping of common and proper nouns. *Word Structure* 4(1). 1–19. doi:10.3366/word.2011.0002

Bibire, Paul. 1998. Moneyers' names on ninth-century Southumbrian coins: Philological approaches to some historical questions. In Mark A. S. Blackburn & D. N. Dumville (eds.), *Kings, currency, and alliances: History and coinage of Southern England in the ninth century*, 155–167. Woodbridge: Boydell & Brewer.

Cannon, Garland. 1989. Abbreviations and acronyms in English word-formation. *American Speech* 64. 99–127. doi:10.2307/455038

Carter, Allyson K. & Cynthia G. Clopper. 2002. Prosodic effects on word reduction. *Language and Speech* 45. 321–353. doi:10.1177/00238309020450040201

Clark, Cecily. 1992. Onomastics. In Richard Hogg (ed.), *The Cambridge history of the English language*, vol. 1, 452–489. Cambridge: Cambridge University Press. doi:10.1017/CHOL9780521264747.008

Coates, Richard. 2006. Names. In Richard Hogg & David Denison (eds.), *A history of the English language*, 312–352. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511791154.007

Colman, Fran. 2014. *The grammar of names in Anglo-Saxon England: The linguistics and culture of the Old English onomasticon*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198701675.001.0001

Davy, Dennis. 2000. Shortening phenomena in Modern English word formation: An analysis of clipping and blending. *Franco-British Studies* 29. 59–76.

Don, Jan. 2014. *Morphological theory and the morphology of English*. Edinburgh: Edinburgh University Press.

Dressler, Wolfgang U. 1997. Universals, typology, and modularity in Natural Morphology. In Raymond Hickey & Stanislaw Puppel (eds.), *Language history and linguistic modelling: A festschrift for Jacek Fisiak on his 60th birthday*, 1399–1424. Berlin: Mouton de Gruyter.

Durkin, Philip. 2009. *The Oxford guide to etymology*. Oxford: Oxford University Press.

Fischer, Roswitha. 1998. *Lexical change in Present-day English*. Tübingen: Gunter Narr.

Hansen, Barbara, Klaus Hansen, Albrecht Neubert & Manfred Schentke. 1990. *Englische Lexikologie: Einführung in Wortbildung und lexikalische Semantik*, 3rd edn. Leipzig: Verlag Enzyklopädie.

Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold.

Hiltunen, Risto. 1983. *The decline of the prefixes and the beginnings of the English phrasal verb: The evidence from some Old and early Middle English texts*. Turku: Turun Yliopisto.

Jamet, Denis. 2009. A morpho-phonological approach of clipping in English: Can the study of clipping be formalized? *Lexis – E-Journal in English Lexicology*. HS1. 15–31. (Special issue 1: Lexicology and phonology.)

Jespersen, Otto. 1922. *Language: Its nature, development and origin*. London: George Allen & Unwin.

Jespersen, Otto. 1933. Monosyllabism in English. *Selected Writings of Otto Jespersen*, 574–598. London: George Allen & Unwin Ltd.

Jespersen, Otto. 1942. *A modern English grammar on historical principles*, vol. 6. London: George Allen & Unwin.

Kager, René. 1999. *Optimality theory*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511812408

Kastovsky, Dieter. 1992. Semantics and vocabulary. In Richard Hogg (ed.), *The Cambridge history of the English language*, vol. 1, 290–408. Cambridge: Cambridge University Press. doi:10.1017/CHOL9780521264747.006

Kastovsky, Dieter. 2006. Vocabulary. In Richard Hogg & David Denison (eds.), *A history of the English language*, 199–270. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511791154.005

Lappe, Sabine. 2007. *English prosodic morphology*. Dordrecht: Springer. doi:10.1007/978-1-4020-6006-9

Marchand, Hans. 1969. *The categories and types of Present-day English word-formation*. München: C.H. Beck.

Mattiello Elisa. 2013. *Extra-grammatical morphology in English: Abbreviations, blends, reduplicatives, and related phenomena*. Berlin: De Gruyter Mouton.

McClure, Peter. 1998. The interpretation of hypocoristic forms of Middle English baptismal names. *Nomina* 21. 101–133.

Miller, Gary D. 2014. *English lexicogenesis*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199689880.001.0001

Minkova, Donka. 1997. Constraint ranking in Middle English stress-shifting. *English Language and Linguistics* 1(1). 135–175. doi:10.1017/S1360674300000393

Minkova, Donka. 2003. *Alliteration and sound change in early English*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511486968

Minkova, Donka. 2008. Prefixation and stress in Old English. *Word Structure* 1(1). 21–52. doi:10.3366/E1750124508000056

Minkova, Donka. 2014. *A historical phonology of English*. Edinburgh University Press.

Minkova, Donka, & Robert P. Stockwell. 2009. *English words: History and structure*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511805929

Nelson, Nicole. 2003. *Asymmetric anchoring*. Amherst: Rutgers University dissertation.

Nevalainen, Terttu. 1999. Early Modern English lexis and semantics. In Roger Lass (ed.), *The Cambridge history of the English language*, vol. 3, 332–458. Cambridge: Cambridge University Press.

Plag, Ingo. 2003. *Word-formation in English*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511841323

Pogatscher, Alois. 1888. *Zur Lautlehre der Griechischen, Lateinischen und Romanischen Lehnworte im Altenglischen*. Strassburg: K.T. Trübner.

Pope, Mildred K. 1934. *From Latin to modern French with especial consideration of Anglo-Norman: Phonology and morphology*. Manchester: Manchester University Press.

Pyles, Thomas & John Algeo. 1993. *The origins and development of the English language*, 4th edn. San Diego: Harcourt Brace Jovanovich.

Ronneberger-Sibold, Elke. 2015. Word-creation. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, 485–499. Berlin: De Gruyter Mouton.

Sundén, Karl. 1904. *Contributions to the study of elliptical words in Modern English*. Uppsala: Almqvist & Wiksells Boktryckeri A.B.

Slettengren, Emrik. 1912. *Contributions to the study of aphæretic words in English*. Lund: Berlingska Boktryckeriet.

Spradlin, Lauren & Taylor Jones. 2016. A morphophonological account of totes constructions in English. Paper presented at the 2016 Annual Meeting of the Linguistic Society of America. Washington, D.C. http://opencuny.org/laurenspradlin/files/2016/01/Totes-LSA-Spradlin-Jones-2016-pdf.pdf (accessed on 9 April, 2018).

Index

The chapters collected in this volume examine how the sociohistorical
and cultural context may influence structural features of lexis and
text types. Each paper pays particular attention to social 'labels' and
attitudes (conservative, religious, ideological, endearing, or other),
thereby focusing on their dynamic and historical dimension. Changes
in these are analyzed in order to explain morphological, lexical,
and textual changes that would otherwise be hard to account for.
Together, they provide a varied window on the effect of historical
versions of a dynamic society on lexis and text. Examining lexical and
textual change in history from a sociocultural perspective teaches us a
great deal – not just about the past, but it also makes us think about
similar phenomena in the present, enhancing our knowledge about
how universally human some of these phenomena are. This volume
will be of great interest to (English) historical linguists, sociolinguists,
and scholars of sociohistorical and cultural studies.

JOHN BENJAMINS PUBLISHING COMPANY